CERTIFICATE OF MAILING
37 C.F.R 1.8

I hereby certify that this correspondence is being deposited with the U.S. Postal Service with sufficient postage as First Class Mail in an envelope addressed to: **MS APPEAL BRIEFS**, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on the date below:

November 17, 2006
Date

David L. Parker

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:
Didier Trono
Maciej Wiznerowicz

Serial No.: 10/720,987

Filed: November 24, 2003

For: COMPOSITIONS AND SYSTEMS FOR THE REGULATION OF GENES

Group Art Unit: 1635

Examiner: Ashen, Jon Benjamin

Atty. Dkt. No.: CLFR:023US

APPEAL BRIEF

MS Appeal Briefs
Commissioner for Patents
P. O. Box 1450
Alexandria, VA 22313-1450

Appellants hereby submit this Appeal Brief to the Board of Patent Appeals and Interferences pursuant to 37 C.F.R. §41.31(a)(1) and 41.37 in light of the Final Office Action dated May 25, 2006. A Notice of Appeal was filed on August 25, 2006. It is believed that the filing of the present Appeal Brief is timely, with one month's extension, and the appropriate fees for filing this Brief are enclosed. However, if any fees are due for any reason relating to the enclosed materials, the Commissioner is authorized to deduct said fees from Fulbright & Jaworski L.L.P. Account No.: 50-1212/CLFR:023US.

11/21/2006 MBELETE1 00000021 10720987
03 FC:2402 250.00 OP

I. REAL PARTY IN INTEREST

The real party in interest is the assignee, Research Development Foundation. The subject matter of this application is licensed to Merck & Co., Inc.

II. RELATED APPEALS AND INTERFERENCES

There are no interferences or appeals for related cases.

III. STATUS OF THE CLAIMS

Claims 1-7, 9-11, 13, 41, 46 and 47 are pending and under examination. Claims 8, 12, 14-40, 42-45 and 48-84 are withdrawn. No claims are canceled. Claims 1-7, 9-11, 13, 41, 46 and 47 are the subject of the present appeal.

IV. STATUS OF AMENDMENTS

There are no pending amendments.

V. SUMMARY OF THE CLAIMED SUBJECT MATTER

The invention, as set forth in independent claim 1, concerns a polynucleotide construct comprising a region encoding a siRNA operably linked to an externally controllable RNA polymerase III promoter, wherein expression of the siRNA is regulated by a polypeptide regulator having both a DNA binding domain and a repressor domain. See specification at page 5, lines 14-21 and page 8, lines 8-22.

The invention as set forth in independent claim 41 concerns a mammalian cell comprising a polynucleotide construct comprising a region encoding a siRNA operably linked to an externally controllable RNA polymerase III promoter, wherein expression of the siRNA is regulated by a polypeptide regulator having both a DNA binding domain and a repressor domain,

said polypeptide regulator being encoded by said cell. Specific examples including undifferentiated cells (claim 46) and oocytes (claim 47). See specification at pages 11 and 12.

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

There are four rejections that are the subject of the present appeal:

1) A rejection of claims 41, 46 and 47 as lacking written description under 35 U.S.C. §112, first paragraph;

2) A rejection of claims 1-7, 9-11, 13, 41 and 46-47 as lacking written description under 35 U.S.C. §112, first paragraph;

3) A rejection of claims 41, 46 and 47 as unpatentable under 35 U.S.C. §101;

4) A rejection of claims 1-7, 9-11, 13, 41 and 46-47 as unpatentable under 35 U.S.C. §103 over Yao *et al.* ("Yao"; Exhibit 1), Verma *et al.* ("Verma"; Exhibit 2), Elbasher *et al.* ("ElbasherA"; Exhibit 3), Elbasher *et al.*/Nature ("ElbasherB"; Exhibit 4) and Deuschle *et al.* ("Deuschle"; Exhibit 5); and

5) A rejection of claims 1-7, 9-11, 13, 41 and 46-47 as unpatentable under 35 U.S.C. §103 over Giordano et al. ("Giordano"; Exhibit 6), ElbasherB and Deuschle and Verma.

VII. ARGUMENT

A. Rejection of claims 41, 46 and 47 as lacking written description under 35 U.S.C. §112, first paragraph

The Final Action first rejected claims 41, 46 and 47 as lacking written description under 35 U.S.C. §112 taking the position that the phrase "wherein expression of the siRNA is regulated by a polypeptide regulator having both a DNA binding domain and a repressor domain, said

polypeptide regulator begin encoded by said cell" constitutes new matter. In particular, the Action posits that the originally filed specification fails to disclose the polypeptide regulator being encoded by a cell. In response, we would note that there are many, many places in the specification where this is discussed. Indeed, it is difficult to visualize carrying out this aspect of the invention without the polypeptide regulator being encoded by the cell.

First, we would direct the Board's attention to excerpts which teach the general concept of expressing poplypeptide regulator, exemplified by tTR-KRAB, by cells to effect controlled expression of siRNA by the cell:

> Expression of tTR-KRAB or other regulatable system may be conditional, inducible, tissue-specific, constitutive, or locally utilized (such as locally applied). A founder cell or cell line can be used for introduction of a nucleic acid sequence containing a sequence of interest under the control of regulatory element for the transcriptional fusion protein (effector polynucleotide), such as tTR-KRAB.

Specification, page 12, lines 19-23.

Next, we would direct the Board's attention to the paragraph bridging pages 20 and 21, wherein the specification describes the preparation of transgenic animals, wherein cells of the transgenic animals are transfected with the coding sequence for the polypeptide regulator:

> Furthermore, such animals may also be regulated in a conditional and tissue-specific manner comprising: a) obtaining a first transgenic animal having a Cre recombinase-encoding polynucleotide under the control of a tissue-specific promoter; b) obtaining a second transgenic animal having i) an siRNA-encoding nucleic acid segment ***under the control of a regulatable promoter region***, wherein the siRNA corresponds to the target gene; and, ***ii) the polynucleotide encoding the regulatable polypeptide regulator***; and, c) mating opposite sexes of the first and second animals. In some embodiments, the second transgenic animal is obtained by: d***) transfecting an undifferentiated mammalian cell with i) a regulatable siRNA-expression construct comprising an siRNA encoding nucleic acid segment and a regulatable promoter region***, wherein an excisable fragment is located between the segment and the regulatable promoter region; and ii) ***a polynucleotide encoding a regulatable polypeptide regulator of the regulatable***

siRNA-expression construct; e) fertilizing the cell if the cell has a haploid genome; and, f) transplanting the embryo into a female animal, wherein the female animal produces the second transgenic animal. In some cases, an undifferentiated cell is an unfertilized oocyte, a fertilized oocyte, an embryonic stem cell, a cell within a morula or blastocyst. Moreover, methods can also involve culturing the cell prior to transfection and/or transplantation. Further, such methods can involve conventional matings with transgenic animals of the invention.

Specification, para. Bridging pages 20-21.

The foregoing concept is also embodied in the originally filed claims:

42. The mammalian cell of claim 41, further defined as comprising:

(a) a first polynucleotide sequence comprising a promoter operably linked to at least one nucleic acid segment encoding an siRNA;

(b) a second polynucleotide sequence encoding a conditional repressor fusion protein that comprises a DNA binding domain and a transcription repression domain; and

(c) a third polynucleotide sequence bindable by the binding domain of the fusion protein of (b) and positioned such that the transcription repression domain acts to repress transcription of the nucleic acid segment of (a).

43. The mammalian cell of claim 41, wherein the conditional repressor fusion protein is drug inducible.

The foregoing claims evidence the fact that the application as originally filed contemplated that the target cells would include and express both the polynucleotide that encoded the siRNA, but also a separate polynucleotide that encoded the fusion protein regulatory element (*i.e.*, the "polypeptide regulator").

B. Rejection of claims 1-7, 9-11, 13, 41 and 46-47 as lacking written description under 35 U.S.C. §112, first paragraph

The Action next rejects claims 1-7, 9-11, 13, 41 and 46-47, taking the position that the concept of the polypeptide regulator having both a DNA binding domain and a repressor domain is insufficiently described in the specification.

In response, we would note that the specification has a detailed description of various examples of DNA binding domains and repressor domains that can be used in the practice of the invention, as can be seem:

> In a specific embodiment, the KRAB domain does not come from Kox-1 but from another zinc-finger protein-containing KRAB domain. Thus, in one embodiment, the KRAB repression domain from the human KOX-1 protein is used as a transcriptional repressor (Thiesen *et al.*, 1990; Margolin *et al.*, 1994; Pengue *et al.*, 1994; Witzgall *et al.*, 1994). In another embodiment, KAP-1, a KRAB co-repressor, is used with KRAB (Friedman *et al.*, 1996), either as part of the same fusion protein or provided separately. Alternatively, KAP-1 can be used alone with a zinc finger protein.
>
> ...
>
> In addition to the exemplary tTR-KRAB fusion protein as an external agent-inducible repressor fusion protein, other fusion proteins comprising different DNA binding domains and transcriptional repressor domains may be utilized. In embodiments of the present invention, a DNA binding domain is utilized as part of a drug-inducible regulatory fusion protein, and the DNA-binding domain may include sequences such as the DNA-binding domains of the tetracycline repressor (tTR), or those of GAL4 or LexA, for example.
>
> In embodiments of the present invention wherein a repressor domain is utilized as part of an external agent-inducible regulatory fusion protein, the repressor domain is the Kruppel-associated box domain (KRAB). However, other repressor domains include ERD or SID transcriptional repressor domains, for example. Other preferred transcription factors and transcription factor domains that act as transcriptional repressors include, for example, MAD (see, *e.g.*, Sommer *et al.*, 1998; Gupta *et al.*, 1998; Queva *et al.*, 1998; Larsson *et al.*, 1997; Laherty *et al.*, 1997; and Cultraro *et al.*, 1997); FKHR (forkhead in rhapdosarcoma gene; Ginsberg *et al.*, 1998; Epstein *et al.*, 1998); EGR-1 (early growth response gene product-1; Yan *et al.*, 1998; and Liu *et al.*, 1998); the ets2

repressor factor repressor domain (ERD; Sgouras et al, 1995); and the MAD smSIN3 interaction domain (SID; Ayer *et al.*, 1996).

Specification, pages 8, 10-11.

From the foregoing, it is evident that the specification sets forth both a large number of known DNA binding domains and separate repressor domains that can be employed in the practice of the invention in order to provide for expression of the polypeptide regulator.

The Action also takes that position that the specification does not adequately describe the "polypeptide regulator." We suspect that the Examiner may be a bit confused as to what is intended by the polypeptide regulator. As noted in the foregoing section, the polypeptide regulator is simply the regulatory fusion protein that includes both the DNA binding domain and the repressor domain. An example is the "tTR-KRAB fusion protein" but as can be seen from the foregoing excerpts, the specification describes many other possible examples.

The Action further complains that the specification fails to adequately describe examples of "externally controllable promoters." Again, we strenuously disagree in that "externally controllable promoters" are exceedingly well known in the art and described in the Summary of specification:

> In a particular embodiment, the invention regards a polynucleotide construct comprising a region encoding a siRNA operably linked to an externally controllable promoter, and the construct may be further defined as a vector. A non-limiting example of a vector includes a lentiviral vector, a retroviral vector, an MLV vector, an AAV vector, a plasmid vector or an adenoviral vector. The externally controllable promoter may be a repressible promoter whereby expression of the encoded siRNA can be downregulated by means of an externally applied agent. The expression of the encoded siRNA can be downregulated by means of an externally applied drug.
>
> In specific embodiments, the repressible promoter is regulated by a Tet repressor and/or is defined as further comprising at least one *tetO* sequence. The repressible promoter may be regulated by the lacI repressor, or the repressible

> promoter may be from the gene of ANB1, HEM 13, ERG 11, OLE 1, GAL1, GAL10, ADH2, or TET^R.
>
> In particular embodiments, the externally controllable promoter is an inducible promoter whereby expression of the encoded siRNA can be upregulated by means of an externally applied agent. Inducible promoters may be inducible by Cu^{2+}, Zn^{2+}, tetracycline, tetracycline analog, ecdysone, glucocorticoid, tamoxifen, or an inducer of the lac operon. The promoter may be inducible by ecdysone, glucocorticoid, or tamoxifen. In specific embodiments, the inducible promoter is a phage inducible promoter, nutrient inducible promoter, temperature inducible promoter, radiation inducible promoter, metal inducible promoter, hormone inducible promoter, steroid inducible promoter, or combination thereof. Examples of radiation inducible promoters include fos promoter, jun promoter, or erg promoter.
>
> Systems for the regulation of gene expression that may be used within the contemplated scope of the invention include regulatory systems utilizing compounds such as progesterone, estrogen, and/or ecdysone.

Specification, pages 5-6.

Furthermore, we would direct the Board to the 14-page section of the specification entitled "Controlled expression of siRNA constructs" at pages 29-42 which includes a very detailed, exemplary listing and description of possible control promoters that could be employed in the practice of the invention.

The foregoing is consistent with Federal Circuit caselaw interpreting the law of written description in the context of biotech inventions such as the present one. Instructive in this regard is the Federal Circuit's recent decision in *Capon v. Eshhar v. Dudas*, 418 F.3d 1349, 76 USPQ2d 1078 (Fed. Cir. 2005). As the *Capon* court points out, there is no requirement under written description that a specification contain a detailed description of elements where those elements are well known to those in the field:

> The Board stated that "controlling precedent" required inclusion in the specification of the complete nucleotide sequence of "at least one" chimeric gene. Bd. op. at 4. The Board also objected that the claims were broader than the specific examples. Eshhar and Capon each responds by pointing to the scientific

completeness and depth of their descriptive texts, as well as to their illustrative examples. The Board did not relate any of the claims, broad or narrow, to the examples, but invalidated all of the claims without analysis of their scope and the relation of claim scope to the details of the specifications.

Eshhar and Capon both argue that they have set forth an invention whose scope is fully and fairly described, for the nucleotide sequences of the DNA in chimeric combination is readily understood to contain the nucleotide sequences of the DNA components. Eshhar points to the general and specific description in his specification of known immune-related DNA segments, including the examples of their linking. Capon points similarly to his description of selecting DNA segments that are known to express immune-related proteins, and stresses the existing knowledge of these segments and their nucleotide sequences, as well as the known procedures for selecting and combining DNA segments, as cited in the specification.

Both parties argue that the Board misconstrued precedent, and that precedent does not establish a *per se* rule requiring nucleotide-by-nucleotide re-analysis when the structure of the component DNA segments is already known, or readily determined by known procedures. The "written description" requirement implements the principle that a patent must describe the technology that is sought to be patented; the requirement serves both to satisfy the inventor's obligation to disclose the technologic knowledge upon which the patent is based, and to demonstrate that the patentee was in possession of the invention that is claimed. See *Enzo Biochem*, 296 F.3d at 1330 (the written description requirement "is the *quid pro quo* of the patent system; the public must receive meaningful disclosure in exchange for being excluded from practicing the invention for a limited period of time"); *Reiffin v. Microsoft Corp.*, 214 F.3d 1342, 1345-46 (Fed. Cir. 2000) (the purpose of the written description requirement "is to ensure that the scope of the right to exclude . . . does not overreach the scope of the inventor's contribution to the field of art as described in the patent specification"); *In re Barker*, 559 F.2d 588, 592 n.4 (CCPA 1977) (the goal of the written description requirement is "to clearly convey the information that an applicant has invented the subject matter which is claimed"). The written description requirement thus satisfies the policy premises of the law, whereby the inventor's technical/scientific advance is added to the body of knowledge, as consideration for the grant of patent exclusivity.

The descriptive text needed to meet these requirements varies with the nature and scope of the invention at issue, and with the scientific and technologic knowledge already in existence. The law must be applied to each invention that enters the patent process, for each patented advance is novel in relation to the state of the science. Since the law is applied to each invention in view of the state of relevant knowledge, its application will vary with differences in the state of knowledge in the field and differences in the predictability of the science.

For the chimeric genes of the Capon and Eshhar inventions, the law must take cognizance of the scientific facts. The Board erred in refusing to consider the state of the scientific knowledge, as explained by both parties, and in declining to

consider the separate scope of each of the claims. None of the cases to which the Board attributes the requirement of total DNA re-analysis, *i.e.*, *Regents v. Lilly*, *Fiers v. Revel*, *Amgen*, or *Enzo Biochem*, require a re-description of what was already known. In *Lilly*, 119 F.3d at 1567, the cDNA for human insulin had never been characterized. Similarly in *Fiers*, 984 F.2d at 1171, much of the DNA sought to be claimed was of unknown structure, whereby this court viewed the breadth of the claims as embracing a "wish" or research "plan." In *Amgen*, 927 F.2d at 1206, the court explained that a novel gene was not adequately characterized by its biological function alone because such a description would represent a mere "wish to know the identity" of the novel material. In *Enzo Biochem*, 296 F.3d at 1326, this court reaffirmed that deposit of a physical sample may replace words when description is beyond present scientific capability. In *Amgen Inc. v. Hoechst Marion Roussel, Inc.*, 314 F.3d 1313, 1332 (Fed. Cir. 2003) the court explained further that the written description requirement may be satisfied "if in the knowledge of the art the disclosed function is sufficiently correlated to a particular, known structure." These evolving principles were applied in *Noelle v. Lederman*, 355 F.3d 1343, 1349 (Fed. Cir. 2004), where the court affirmed that the human antibody there at issue was not adequately described by the structure and function of the mouse antigen; and in *University of Rochester v. G.D. Searle & Co.*, 358 F.3d 916, 925-26 (Fed. Cir. 2004), where the court affirmed that the description of the COX-2 enzyme did not serve to describe unknown compounds capable of selectively inhibiting the enzyme.

The "written description" requirement must be applied in the context of the particular invention and the state of the knowledge. The Board's rule that the nucleotide sequences of the chimeric genes must be fully presented, although the nucleotide sequences of the component DNA are known, is an inappropriate generalization. When the prior art includes the nucleotide information, precedent does not set a *per se* rule that the information must be determined afresh. Both parties state that a person experienced in the field of this invention would know that these known DNA segments would retain their DNA sequences when linked by known methods. Both parties explain that their invention is not in discovering which DNA segments are related to the immune response, for that is in the prior art, but in the novel combination of the DNA segments to achieve a novel result.

The "written description" requirement states that the patentee must describe the invention; it does not state that every invention must be described in the same way. As each field evolves, the balance also evolves between what is known and what is added by each inventive contribution. Both Eshhar and Capon explain that this invention does not concern the discovery of gene function or structure, as in Lilly. The chimeric genes here at issue are prepared from known DNA sequences of known function. The Board's requirement that these sequences must be analyzed and reported in the specification does not add descriptive substance. The Board erred in holding that the specifications do not meet the written description requirement because they do not reiterate the structure or formula or chemical name for the nucleotide sequences of the claimed chimeric genes.

Applicants highlight a quote from the preceding passage: "None of the cases to which the Board attributes the requirement of total DNA re-analysis, *i.e.*, *Regents v. Lilly*, *Fiers v. Revel*, *Amgen*, or *Enzo Biochem*, require a re-description of what was already known." While *Capon* dealt with DNA sequences *per se*, we submit that the reasoning is fully consistent with the present case, which concerns genetic control elements.

The Capon case has very recently been followed by the Federal Circuit, in *Falkner v. Inglis*, App. No. 05-1324, decided May 26, 2006 (Fed. Cir. 2006) (copy enclosed). In a section of the opinion entitled "Recitation of Known Structure Is Not Required" the *Falkner* court, following the *Capon* decision, stated:

> Indeed, a requirement that patentees recite known DNA structures, if one existed, would serve no goal of the written description requirement. It would neither enforce the quid pro quo between the patentee and the public by forcing the disclosure of new information, nor would it be necessary to demonstrate to a person of ordinary skill in the art that the patentee was in possession of the claimed invention. ... Accordingly, we hold that where, as in this case, accessible literature sources clearly provided, as of the relevant date, genes and their nucleotide sequences (here "essential genes"), satisfaction of the written description requirement does not require either the recitation or incorporation by reference (where permitted) of such genes and sequences.

Id. at page 17-18. While *Falkner* dealt with nucleotide and gene sequences, the reasoning is certainly applicable here as well.

Thus, it is submitted that there is clearly no *prima facie* basis for the subject written description rejection.

C. Rejection of claims 41, 46 and 47 as unpatentable under 35 U.S.C. §101.

The Action next rejects claims 41, 46 and 47 as directed to non-statutory subject matter in that the claims are said to read on transgenic humans.

In response, Applicants would note that while the claim could potentially cover some future transgenic human, the claims are not "directed to" a transgenic human and thus is not "directed to" non-statutory subject matter.

Following the Examiner's logic, a claim directed to a transgenic DNA sequence could also be read to cover some hypothetical transgenic human, and yet such claims are routinely issued as are claims to transgenic mammalian cells.

Further with respect to claims 46 (undifferentiated cells) and 47 (oocytes), these claims cannot be interpreted to read on a whole human.

D. Rejection of claims 1-7, 9-11, 13, 41 and 46-47 as unpatentable under 35 U.S.C. §103 over Yao, Verma, ElbasherA, ElbasherB and Deuschle.

Next, the Action rejects all of the pending and examined claims as unpatentable under 35 U.S.C. §103 over Yao (Exhibit 1), Verma (Exhibit 2), ElbasherA (Exhibit 3), ElbasherB (Exhibit 4) and Deuschle (Exhibit 5).

Appellants note that Verma teaches lentiviruses in general, and their use in expressing heterologous nucleic acids, but observe that Verma is admittedly silent with respect to the expression of siRNA expression constructs operably linked to an externally controllable RNA polymerase III promoter, wherein expression of the siRNA is regulated by a polypeptide regulator having both a DNA binding domain and a repressor domain.

The Yao reference teaches nucleic acid constructs, including retroviruses, for expressing heterologous sequences under the control of a Tet repressor, but is also silent with respect to the expression of siRNA expression constructs operably linked to an externally controllable RNA polymerase III promoter, wherein expression of the siRNA is regulated by a polypeptide regulator having both a DNA binding domain and a repressor domain.

The two Elbashir references relate generally to siRNA expression constructs, but do not appear to in any way teach or suggest their being operably linked to an externally controllable RNA polymerase III promoter, wherein expression of the siRNA is regulated by a polypeptide regulator having both a DNA binding domain and a repressor domain. Indeed, we have been unable to identify any teaching or suggestion to control siRNA expression through the use of any externally controllable promoter, much less one that has both a DNA binding domain and a repressor domain.

Lastly, the Deuschle reference discloses an externally controllable promoter having a DNA binding domain and a repressor domain (TetR-KRAB), but we have been unable to find any teaching here relevant to the use of such a promoter in the context of an siRNA under the control of a pol III promoter as specified in the claims. For example, the TetR-KRAB constructs is employed here only in the context of the CMV promoter, which is a pol II not a pol III promoter. Further, the only heterologous gene that Appellant's have identified in Deuschle that is placed under the control of TetR-KRAB is the firefly luciferase gene.

Thus, even if the references are properly combinable (which, as discussed below, we dispute) they nevertheless fail to teach or suggest the invention as claimed in claim 1, in that the combination fails to teach a polynucleotide construct comprising a region encoding a siRNA operably linked to an externally controllable RNA polymerase III promoter, wherein expression of the siRNA is regulated by a polypeptide regulator having both a DNA binding domain and a repressor domain as required by independent claims 1 and 41.

However, we further submit that the Examiner has failed to demonstrate that the references are properly combinable. It appears that the only references relied upon that concern

siRNA, the Elbasher references, fail to teach the concept of controlled expression of siRNA. Indeed, we have been unable to identify any such teaching in any of the references relied upon by the Examiner. Accordingly, we fail to understand the Examiner's *prima facie* basis for combining the Elbasher references with the other references as proposed. The same can be said for the other references. Deuschle teaches a CMV promoter that is regulated by TetR-KRAB, but the Examiner has failed to explain how this reference is combinable with references that concern other required aspects of the claims, such as siRNA expression. Although not entirely clear from the record, the Examiner may be attempting to rely on the Verma's disclosure of antisense and ribozyme expression (col. 7). The same can be said for Yao's teaching of controlled expression of antisense molecules (page 20 "D"). However, antisense and ribozymes are dramatically different from siRNAs, which is seen most notably in the fact that the latter is expressed from pol III promoters, not pol II promoters, and in the fact that their underlying mechanisms of action are quite distinct.

Here, the Examiner has failed to enunciate any motivation to combine the references in the manner suggested, and has failed, for example, to make a showing that it is within the "common knowledge and common sense" to expand the cited references to include the concept of using a pol III promoter in this construct. *Dystar Textilfarben Gmbh v. C. H. Patrick Co.*, 2006 U.S. App. LEXIS 24642 (Fed. Cir. 2006). The Court of Appeals for the Federal Circuit stated this concept succinctly in *In re Kotzab*:

> ... Even when obviousness is based on a single prior art reference, there must be a showing of a suggestion or motivation to modify the teachings of that reference. See B. F. Goodrich Co. v. Aircraft Braking Sys. Corp., 72 F.3d 1577, 1582, 37 U.S.P.Q. 2d 1314, 1318 (Fed. Cir. 1996).
>
> The motivation, suggestion or teaching may come explicitly from statements in the prior art, the knowledge of one of ordinary skill in the art, or, in some cases the nature of the problem to be solved. See Dembiczak, 175 F.3d at

> 999, 50 U.S.P.Q. 2D (BNA) at 1617. In addition, the teaching, motivation or suggestion may be implicit from the prior art as a whole, rather than expressly stated in the references. See WMS Gaming, Inc. v. International Game Tech., 184 F.3d 1339, 1355, 51 U.S.P.Q. 2D (BNA) 1385, 1397 (Fed. Cir. 1999). The test for an implicit showing is what the combined teachings, knowledge of one of ordinary skill in the art, and the nature of the problem to be solved as a whole would have suggested to those of ordinary skill in the art. See In re Keller, 642 F.2d 413, 425, 208 U.S.P.Q. (BNA) 871, 881 (CCPA 1981) (and cases cited therein). Whether the Board relies on an express or an implicit showing, it must provide particular findings related thereto. See Dembiczak, 175 F.3d at 999, 50 U.S.P.Q. 2D (BNA) at 1617. Broad conclusory statements standing alone are not "evidence." Id.

In re Kotzab, 217 F.3d. 1365, 1370, 55 U.S.P.Q. 2d 1313, 1317 (emphasis in original).

Thus, to make out a proper *prima facie* obviousness rejection, the Examiner must adequately ***explain*** the motivation to combine separate teachings directed to 1) externally regulatable siRNA expression, 2) under the control of an RNA polymerase III promoter, 3) where the siRNA expression is regulated by a polypeptide regulator having both a DNA binding domain and a repressor domain (such as TetR/KRAB). See MPEP §2143.01 The mere fact that references "can be combined or modified" or that the claimed invention is "within the capabilities of one of ordinary skill" is clearly insufficient. *Id.* at sections III and IV. Furthermore, the Examiner must demonstrate on the record that one of ordinary skill had a reasonable expectation that such a combination would be successful. MPEP §2143.02 ("Reasonable expectation of success required"). Neither of the foregoing has been set forth by the Examiner.

Claims 46 and 47 Are Separately Patentable

The subject matter of claims 46 and 47 are even further removed from the prior art.

Claim 46 is directed to an undifferentiated mammalian cell that comprises a polynucleotide construct comprising a region encoding a siRNA operably linked to an externally controllable RNA polymerase III promoter, wherein expression of the siRNA is regulated by a

polypeptide regulator having both a DNA binding domain and a repressor domain, said polypeptide regulator being encoded by said cell. The Examiner apparently relies on Yao for such a suggestion. Action at 8 (citing page 15 of Yao). However, the section of Yao relied upon by the Examiner is very clearly only concerned with a "Method for Recombinantly Producing ***Protein***" (emphasis ours). Thus, we see no basis in Yao (or any of the other references) for any suggestion to employ externally controlled siRNA in the context of undifferentiated cells such as stem cells.

Claim 47 is directed to oocytes or fertilized oocytes that comprise a externally controllable siRNA expression construct and are even still further removed. Applicant's have been unable to identify what teaching, if any, is being relied upon with respect to the subject matter of claim 47.

C. A rejection of claims 1-7, 9-11, 13, 41 and 46-47 as unpatentable under 35 U.S.C. §103 over Giordano et al. ("Giordano"; Exhibit 6), ElbasherB and Deuschle and Verma.

Lastly, the Action rejects claims 1-7, 9-11, 13, 41 and 46-47 over many of the same references, but with the addition of the Giordano reference. The Action indicates that Giordano is cited as teaching inducible and repressible transcription systems that can be used to control the timing of the expression of dsRNA from, for example, retroviral vectors, and that an example of such a construct is one incorporating the Tet promoter. Action at pages 12-13.

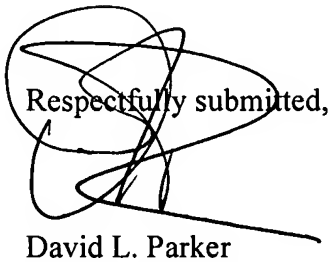
With respect to all of the rejections except Giordano, we incorporate the arguments set forth in the preceding section. With respect to Giordano, we have reviewed this reference and have been unable to find any teaching or suggestion of an siRNA operably linked to an externally controllable RNA polymerase III promoter, wherein expression of the siRNA is regulated by a polypeptide regulator having both a DNA binding domain and a repressor domain.

While it does appear to teach a DNA binding domain, such as is found in the TetR, we can not find any suggestion to use a polypeptide regulator having ***both*** a DNA binding domain ***and*** a repressor domain such as TetR/KRAB. Furthermore, this reference teaches away from using such a polypeptide regulator for inhibiting RNA polymerase III. ("...these factors carry protein domains that transactivate or transrepress ***the RNA polymerase II***." Col 33, line 19-21 (emphasis ours)).

For the foregoing reasons, the Examiner has still failed to make out a proper *prima facie* obviousness rejection for the same reasons as set forth in the previous section of this brief.

VIII. CONCLUSION

Appellants believe that the foregoing remarks fully respond to all outstanding matters for this application. Appellants respectfully request that the Board reverse the rejections of all claims.



Respectfully submitted,

David L. Parker
Reg. No. 32,165
Attorney for Applicants

FULBRIGHT & JAWORSKI L.L.P.
600 Congress Avenue, Suite 2400
Austin, Texas 78701
(512) 474-5201
(512) 536-4598 (facsimile)

Date: November 17, 2006

I. CLAIMS APPENDIX

1. (Previously presented) A polynucleotide construct comprising a region encoding a siRNA operably linked to an externally controllable RNA polymerase III promoter, wherein expression of the siRNA is regulated by a polypeptide regulator having both a DNA binding domain and a repressor domain.

2. (Original) The construct of claim 1, further defined as a vector.

3. (Original) The construct of claim 2, wherein the vector is a lentiviral vector, a retroviral vector, an MLV vector, an AAV vector, a plasmid vector or an adenoviral vector.

4. (Original) The construct of claim 1, wherein the externally controllable promoter is a repressible promoter whereby expression of the encoded siRNA can be downregulated by means of an externally applied agent.

5. (Original) The construct of claim 4, wherein expression of the encoded siRNA can be downregulated by means of an externally applied drug.

6. (Previously presented) The construct of claim 1, wherein the externally controllable promoter is a repressible promoter that is regulated by a Tet repressor which further comprises a DNA binding domain.

7. (Previously presented) The construct of claim 1, wherein the externally controllable promoter is a repressible promoter comprising at least one *tetO* sequence.

8. (Withdrawn) The construct of claim 1, wherein the repressible promoter is regulated by the lacI repressor.

9. (Previously presented) The construct of claim 1, wherein the externally controllable promoter is a repressible promoter from the gene of ANB1, HEM 13, ERG 11, OLE 1, GAL1, GAL10, ADH2, or TET^R.

10. (Original) The construct of claim 1, wherein the externally controllable promoter is an inducible promoter whereby expression of the encoded siRNA can be upregulated by means of an externally applied agent.

11. (Original) The construct of claim 10, wherein the inducible promoter is inducible by Cu^{+2}, Zn^{2+}, tetracycline, tetracycline analog, ecdysone, glucocorticoid, tamoxifen, or an inducer of the lac operon.

12. (Withdrawn) The construct of claim 11, wherein said promoter is inducible by ecdysone, glucocorticoid, or tamoxifen.

13. (Original) The construct of claim 10, wherein said inducible promoter is a phage inducible promoter, nutrient inducible promoter, temperature inducible promoter, radiation

inducible promoter, metal inducible promoter, hormone inducible promoter, steroid inducible promoter, antibiotic inducible promoter, or combination thereof.

14. (Withdrawn) The construct of claim 13, wherein said radiation inducible promoter is a fos promoter, a jun promoter, or an erg promoter.

15. (Withdrawn) A system for controlling gene expression comprising:

(a) one or more polynucleotide constructs in accordance with claim 1;

(b) a compound or reagent that may be administered to the cell that directly or indirectly controls the expression of the siRNA.

16. (Withdrawn) The system of claim 15, further defined as comprising:

a) a regulatable siRNA-expression construct comprising an siRNA encoding nucleic acid segment potentially under the control of a regulatable promoter region; and

b) a polynucleotide encoding a regulatable polypeptide regulator of the regulatable siRNA-expression construct.

17. (Withdrawn) The system of claim 15, wherein an excisable fragment is located between the segment and the regulatable promoter region.

18. (Withdrawn) The system of claim 17, wherein the expression of an agent that excises said excisable fragment is under the control of a tissue-specific promoter.

19. (Withdrawn) The system of claim 16, wherein the regulatable siRNA-expression construct is an integrating viral vector.

20. (Withdrawn) The system of claim 18, wherein the integrating viral vector is a lentivirus, a retrovirus, or an adeno-associated virus.

21. (Withdrawn) The system of claim 19, wherein the integrating viral vector is a lentivirus.

22. (Withdrawn) The system of claim 16, wherein the regulatable siRNA-expression construct comprises a marker to assay for integration.

23. (Withdrawn) The system of claim 16, wherein the regulatable promoter region comprises a nucleic acid segment that negatively regulates transcription from the promoter.

24. (Withdrawn) The system of claim 16, wherein the nucleic acid segment that negatively regulates transcription from the promoter is a *tet* operator.

25. (Withdrawn) The system of claim 22, wherein the regulatable polypeptide regulator is capable of binding the *tet* operator.

26. (Withdrawn) The system of claim 18, wherein the tissue specific promoter is liver fatty acid binding (FAB) protein gene promoter, insulin gene promoter, transphyretin promoter, α1-antitrypsin promoter, plasminogen activator inhibitor type 1 (PAI-1) promoter, apolipoprotein AI promoter, LDL receptor gene promoter, myelin basic protein (MBP) gene promoter, glial

fibrillary acidic protein (GFAP) gene promoter, opsin promoter, LCK promoter, CD4 promoter, keratin promoter, myoglobulin promoter, or neural-specific enolase (NSE) promoter.

27. (Withdrawn) The system of claim 16, wherein the polynucleotide is under the control of a constitutive promoter.

28. (Withdrawn) The system of claim 16, wherein the polynucleotide is comprised in a viral vector.

29. (Withdrawn) The system of claim 27, wherein the viral vector is an integrating viral vector.

30. (Withdrawn) The system of claim 28, wherein the integrating viral vector is a lentivirus, retrovirus, or adeno-associated virus.

31. (Withdrawn) The system of claim 16, further defined as comprising:

(a) a polynucleotide construct comprising a promoter operably linked to at least one polynucleotide encoding siRNAs;

(b) a polynucleotide encoding an inducible repressor that can repress the expression of said at least one siRNA;

(c) one or more vectors comprising the constructs of (a) and (b); and

(e) a compound or reagent that may be administered to the cell that controls the expression of the fusion protein.

32. (Withdrawn) The system of claim 31, further defined as comprising:

(a) a polynucleotide construct comprising a polymerase III-dependent promoter operably linked to at least one polynucleotide encoding siRNAs;

(b) a polynucleotide encoding a drug-inducible repressor fusion protein that comprises a DNA binding domain and a transcription repression domain; and

(c) a polynucleotide bindable by the binding domain of the fusion protein of (b) and positioned such that the transcription repression domain acts to repress transcription of the polynucleotide construct of (a);

(d) one or more vectors comprising the constructs of (a), (b), and (c); and

(e) a compound that may be administered to the cell that controls the expression of the fusion protein or that controls the binding of the fusion protein to the polynucleotide sequence bindable by the binding domain of the fusion protein.

33. (Withdrawn) The system of claim 32, wherein the polynucleotide encoding the fusion protein is operatively linked to an inducible promoter.

34. (Withdrawn) The system of claim 32, wherein the promoter is a constitutive promoter.

35. (Withdrawn) The system of claim 32, wherein the polynucleotide sequence bindable by the binding domain of the fusion protein is the tetracycline operator (*tetO*) sequence, the fusion protein of (ii) is comprised of the DNA binding domain of the tetracycline repressor (tTR) fused to the KRAB repression domain of human Kox-1 (tTR-KRAB), and the substance of (c) is doxycycline.

36. (Withdrawn) The system of claim 32, wherein the vector of (b) is a lentiviral vector, an MLV vector, an AAV vector, a plasmid vector or an adenoviral (Adv or Ad) vector.

37. (Withdrawn) The system of claim 35, wherein the polynucleotide sequence bindable by the binding domain of the fusion protein of (b), the promoter operably linked to the polynucleotide sequence encoding the siRNA and the polynucleotide sequence encoding the siRNA are comprised in the U3 region of the 3' long terminal repeat of the lentiviral vector.

38. (Withdrawn) The system of claim 32, wherein the polynucleotide encoding the fusion protein is comprised within a second, separate vector from the vector comprising the constructs of (a).

39. (Withdrawn) The system of claim 38, wherein the second vector comprising the polynucleotide encoding the fusion protein is a lentiviral vector, a MLV vector, an AAV vector, a plasmid vector or an adenoviral (Adv or Ad) vector.

40. (Withdrawn) The system of claim 15, wherein the polynucleotide encodes siRNA that forms a stem-and-loop structure, or a hairpin.

41. (Previously presented) A mammalian cell comprising a polynucleotide construct comprising a region encoding a siRNA operably linked to an externally controllable RNA polymerase III promoter, wherein expression of the siRNA is regulated by a polypeptide regulator having both a DNA binding domain and a repressor domain, said polypeptide regulator being encoded by said cell.

42. (Withdrawn) The mammalian cell of claim 41, further defined as comprising:

(a) a first polynucleotide sequence comprising a promoter operably linked to at least one nucleic acid segment encoding an siRNA;

(b) a second polynucleotide sequence encoding a conditional repressor fusion protein that comprises a DNA binding domain and a transcription repression domain; and

(c) a third polynucleotide sequence bindable by the binding domain of the fusion protein of (b) and positioned such that the transcription repression domain acts to repress transcription of the nucleic acid segment of (a).

43. (Withdrawn) The mammalian cell of claim 41, wherein the conditional repressor fusion protein is drug inducible.

44. (Withdrawn) The mammalian cell of claim 42, further comprising

(d) a fourth polynucleotide sequence, wherein the fourth polynucleotide sequence is excisable and prevents transcription from the polymerase III-dependent promoter; and,

(e) a fifth polynucleotide sequence encoding an enzyme capable of excising the fourth polynucleotide sequence, wherein the fifth polynucleotide sequence is under the control of a regulatable promoter.

45. (Withdrawn) The mammalian cell of claim 44, wherein the third polynucleotide sequence bindable by the binding domain of the fusion protein is the tetracycline operator (*tetO*) sequence, the fusion protein of (b) is comprised of the DNA binding domain of the tetracycline repressor (tTR) fused to the KRAB repression domain of human Kox-1 (tTR-KRAB), and the fusion protein is controlled by doxycycline.

46. (Original) The mammalian cell of claim 41, wherein the cell is an undifferentiated cell.

47. (Original) The mammalian cell of claim 41, wherein the cell is an oocyte or fertilized oocyte.

48. (Withdrawn) A transgenic animal having one or more cells as defined by any one of claim 41.

49. (Withdrawn) A method of creating a transgenic animal capable of exhibiting conditional knockdown of a target gene comprising:

a) introducing into a sex cell or an undifferentiated embryonic cell an expression construct comprising a polynucleotide in accordance with claim 1;

b) fertilizing the cell to create an embryo if the cell is a sex cell; and,

c) transplanting the embryo into a female animal, wherein the female animal produces a transgenic animal.

50. (Withdrawn) The method of claim 49, wherein the sex cell or undifferentiated embryonic cell is an unfertilized oocyte, a fertilized oocyte, an embryonic stem cell, a cell within a morula or blastocyst.

51. (Withdrawn) The method of claim 49, further comprising culturing the cell prior to introduction of the expression construct and/or transplantation.

52. (Withdrawn) A method of regulating the expression of a gene in a cell, the method comprising the steps of:

(a) preparing a polynucleotide construct in accordance with claim 1, wherein the siRNA encoded by said construct downregulates the expression of said gene;

(b) externally regulating the expression of the encoded siRNA through the externally controllable promoter.

53. (Withdrawn) The method of claim 52, wherein the externally controllable promoter is a repressible promoter whereby expression of the encoded siRNA can be downregulated by means of an externally applied agent.

54. (Withdrawn) The method of claim 53, wherein expression of the encoded siRNA can be downregulated by means of an externally applied drug.

55. (Withdrawn) The method of claim 53, wherein the repressible promoter comprises at least one *tetO* sequence.

56. (Withdrawn) The method of claim 52, wherein said method further comprises the step of providing a polynucleotide encoding an inducible repressor that can repress the expression of the siRNA.

57. (Withdrawn) The method of claim 56, wherein the polynucleotide encoding the repressor is further defined as a drug-controllable repressor fusion protein that comprises a DNA binding domain and a transcription repression domain, and wherein the binding domain of the fusion protein can bind the polynucleotide construct such that the trancription repression domain acts to repress transcription of the siRNA.

58. (Withdrawn) The method of claim 52, wherein the promoter regulating expression of the siRNA is a constitutive promoter.

59. (Withdrawn) The method of claim 52, wherein the promoter regulating expression of the siRNA is a tissue-specific promoter.

60. (Withdrawn) The method of claim 52, wherein said cell is further defined as being in a cell line.

61. (Withdrawn) The method of claim 59, wherein said method further comprises the step of providing to said cell a drug for testing drug function in the absence of a gene product encoded by said gene.

62. (Withdrawn) The method of claim 52, wherein said cell is comprised in an animal.

63. (Withdrawn) The method of claim 56, wherein the polynucleotide construct encoding the siRNA is further defined as being administered to a region of said animal, wherein the animal comprises said polynucleotide encoding the repressor.

64. (Withdrawn) The method of claim 63, wherein said administering is by injection.

65. (Withdrawn) The method of claim 63, wherein said region of the animal is in an organ.

66. (Withdrawn) The method of claim 62, wherein the animal is a human patient.

67. (Withdrawn) The method of claim 66, wherein the gene is an oncogene.

68. (Withdrawn) The method of claim 67, wherein the oncogene is ABLI, BLC1, BCL6, CBFA1, CBL, CSFIR, ERBA, ERBB, EBRB2, ETS1, ETS1, ETV6, FGR, FOX, FYN, HCR, HRAS, JUN, KRAS, LCK, LYN, MDM2, MLL, MYB, MYC, MYCL1, MYCN, NRAS, PIM1, PML, RET, SRC, TAL1, TCL3 or YES.

69. (Withdrawn) The method of claim 52, wherein externally regulating the expression of the encoded siRNA comprises administering an agent to the patient that effects an upregulation or downregulation of said expression.

70. (Withdrawn) The method of claim 69, wherein the agent is an antibiotic, radiation, steroid, hormone, metal, divalent cation, nutrient, or temperature.

71. (Withdrawn) The method of claim 70, wherein the radiation is ionizing radiation.

72. (Withdrawn) A method of controlling the ability of a cell to be recognized immunologically, comprising the steps of:

(a) obtaining a cell comprising:

(i) a polynucleotide construct in accordance with claim 1, wherein the siRNA encoded by said construct downregulates the expression of a polynucleotide encoding a transplantation antigen; and

(ii) a polynucleotide encoding an inducible repressor that can repress the expression of said siRNA; and

(b) externally regulating the expression of the encoded siRNA through the externally controllable promoter.

73. (Withdrawn) The method of claim 72, wherein said cell is in an animal.

74. (Withdrawn) The method of claim 72, wherein the transplantation antigen is an MHC I antigen.

75. (Withdrawn) The method of claim 72, wherein said transplantation antigen is beta2-microglobulin, an HLA, H-Y, P35B, Kdm4, Kdm5, TL, P198, P91A, or H-2Kb.

76. (Withdrawn) The method of claim 75, wherein said HLA antigen is HLA-C, HLA-G, or HLA-DQ.

77. (Withdrawn) The method of claim 74, wherein said cell is an islet cell, a stem cell, a hepatocyte, a dopaminergic neuron, or a keratinocyte.

78. (Withdrawn) A method of treating a disease condition amenable to treatment with an siRNA, the method comprising the steps of:

(a) preparing a polynucleotide construct in accordance with claim 1, wherein the siRNA encoded by said construct is for the treatment of the disease condition;

(b) externally regulating the expression of said siRNA through said externally controllable promoter.

79. (Withdrawn) The method of claim 78, wherein the disease is a hyperproliferative disorder.

80. (Withdrawn) The method of claim 79, wherein the hyperproliferative disorder is cancer.

81. (Withdrawn) The method of claim 80, wherein the cancer is gliosarcoma, breast cancer, lung cancer, brain cancer, melanoma, prostate cancer, ovarian cancer, pancreatic cancer, liver cancer, colon cancer, cervical cancer, bladder cancer, spleen cancer, head and neck cancer, or bone cancer.

82. (Withdrawn) The method of claim 78, wherein the disease condition is hyperthyroidism.

83. (Withdrawn) The method of claim 78, wherein said disease condition is associated with a hypersecretion defect.

84. (Withdrawn) The method of claim 83, wherein the hypersecretion defect comprises a hormone hypersecretion defect.

II. EVIDENCE APPENDIX

Exhibit 1 -- Yao *et al.*, Office Action dated 12/12/05

Exhibit 2 -- Verma *et al.*, Office Action dated 12/12/05

Exhibit 3 -- Elbasher *et al.*, Office Action dated 12/12/05

Exhibit 4 -- Elbasher *et al.*/Nature, Office Action dated 12/12/05

Exhibit 5 -- Deuschle *et al.*, Office Action dated 5/25/06

Exhibit 6 -- Giordano et al., Office Action dated 12/12/05

EXHIBIT 1

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/85, 15/86, A01K 67/027, A61K 48/00 | A1

(11) International Publication Number: WO 99/00510

(43) International Publication Date: 7 January 1999 (07.01.99)

(21) International Application Number: PCT/US98/10907

(22) International Filing Date: 29 May 1998 (29.05.98)

(30) Priority Data:
08/883,327 26 June 1997 (26.06.97) US

(71) Applicant: BRIGHAM AND WOMEN'S HOSPITAL [US/US]; 1249 Boylston Street, Boston, MA 02215 (US).

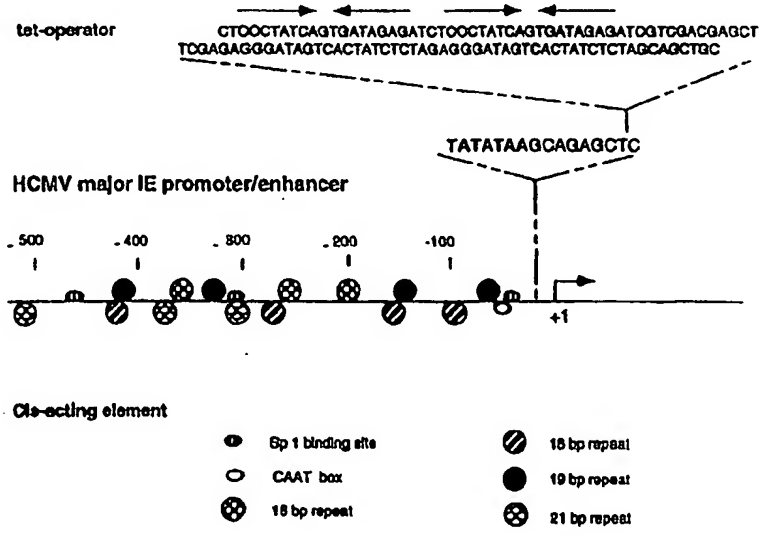
(72) Inventor: YAO, Feng; Apartment #8, 410 Langley Road, Newton Center, MA 02159 (US).

(74) Agent: SANZO, Michael, A.; Vinson & Elkins L.L.P., 2300 First City Tower, 1001 Fannin Street, Houston, TX 77002-6760 (US).

(81) Designated States: CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: REGULATION OF TRANSCRIPTION IN MAMMALIAN CELLS AND VIRAL REPLICATION BY A TETRACYCLIN REPRESSOR



(57) Abstract

The present invention is directed to DNA constructs suitable for gene expression in mammalian cells and which are characterized by the presence of a mammalian promoter under the control of a tet operator/repressor system. The DNA may be used as part of a system for expressing recombinant protein. In addition, the tet operator/repressor system can be used to engineer cis- and trans-destructive viruses which are capable of replicating in the presence of the tet repressor, but not in the absence of the repressor. These viruses can be used either directly in the treatment of patients with corresponding viral diseases, as vehicles for the delivery of nucleic acids that can serve as therapeutic agents and as part of vaccines designed to immunize people or animals against viral diseases.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

REGULATION OF TRANSCRIPTION IN MAMMALIAN CELLS AND VIRAL REPLICATION BY A TETRACYCLIN REPRESSOR

Field of the Invention

The present invention is concerned with compositions and methods that rely upon the tetracycline resistance (tet) operator and repressor to control transcription in mammalian cells. It encompasses methods for recombinantly producing proteins and the vectors and host cells utilized in such methods. In addition, the present invention is directed to viruses which are recombinantly engineered so that their replication is controlled by the tet operator/repressor system. These viruses may serve as vehicles for gene transfer both *in vitro* and *in vivo*; as agents for immunization; and as a means for delivering nucleic acid therapeutic agents to cells.

Background of the Invention

The ability to specifically regulate transgene expression has been a central concern in molecular biology for many years. In the case of mammalian cells, the *in vitro* regulation of recombinant genes has most often been accomplished through the use of inducible promoters that respond to agents such as heavy metal ions (Brinster, *et al.*, *Nature 296*:39-42 (1982); heat shock (Nover, in Heat Shock Response, pp. 167-220, CRC, Fla. (1991)); and hormones (Klock, *et al.*, *Nature 329*:734-736 (1987)). Unfortunately, these promoters generally provide only a relatively a low level of expression even in the presence of inducer and most of the inducers that have been used *in vitro* have unacceptable side effects *in vivo*.

As an alternative to inducible promoters, attempts have been made to control mammalian gene expression using well-characterized prokaryotic regulatory elements. In most cases, regulatory systems have relied upon strong interactions between prokaryotic operators and repressor proteins as a means for either targeting eukaryotic transcription modulators to specific sites within a host cell genome *(see e.g.*, Labow, *et al.*, *Mol. Cell. Biol. 10*:3343-3356 (1990)) or in attempts to directly inhibit gene expression using the prokaryotic repressor (*see e.g.*, Brown, *et al.*, *Cell 49*:603-612 (1987)).

In the case of prokaryotic elements associated with the tetracycline resistance (tet) operon, systems have been developed in which the tet repressor protein is fused with polypeptides known to modulate transcription in mammalian cells. The fusion protein has then been directed

to specific sites by the positioning of the tet operator sequence. For example, the tet repressor has been fused to a transactivator (VP16) and targeted to a tet operator sequence positioned upstream from the promoter of a selected gene (Gussen, *et al., Proc. Nat'l Acad. Sci. USA 89*:5547-5551 (1992); Kim, *et al., J. Virol. 69*:2565-2573 (1995); Hennighausen, *et al., J. Cell. Biochem. 59*:463-472 (1995)). The tet repressor portion of the fusion protein binds to the operator thereby targeting the VP16 activator to the specific site where the induction of transcription is desired. An alternative approach has been to fuse the tet repressor to the KRAB repressor domain and target this protein to an operator placed several hundred base pairs upstream of a gene. Using this system, it has been found that the chimeric protein, but not the tet repressor alone, is capable of producing a 10 to 15-fold suppression of CMV-regulated gene expression (Deuschle, *et al., Mol. Cell. Biol. 15*:1907-1914 (1995)). The main problem with these types of systems is that the portion of fusion proteins corresponding to the mammalian transactivator or repressor tends to interact with cellular transcriptional factors and cause pleiotropic effects.

Ideally, a system for regulating mammalian gene expression should be highly specific for a selected gene and subject to induction by factors suitable for use both *in vitro* and *in vivo*. The present invention discloses such a system and describes how it can be used to regulate transgene expression. In addition, the invention describes how this system can be adapted to engineer viruses to serve as vectors, therapeutic agents and vaccines.

Summary of the Invention

The present invention is directed to a number of different compositions and methods which share the common feature of having gene expression regulated by the tet operator/repressor system.

A. Compositions and Methods for the Production of Recombinant Protein

In its first aspect, the invention is directed to a recombinant DNA molecule which contains a mammalian promoter sequence with a TATA element; at least one tet operator sequence; and a gene sequence operably linked to the promoter and lying downstream from the operator. The exact positioning of the operator sequence (or sequences) relative to the TATA element is

critical to the invention. In order to be effective at controlling transcription, the operator must begin at least 6 nucleotides downstream from the last nucleotide in the TATA element and, when a gene encoding a protein is expressed, the operator should be positioned before the translation initiation codon. In general, the operator should not begin more than about 100 nucleotides downstream and, preferably, it should begin within 6 to 24 nucleotides downstream of the TATA element. When positioned in this manner, it has been found that the binding of the repressor protein causes an essentially complete shutdown in transcriptional activity. This is true even for very strong and highly promiscuous promoters such as the human CMV immediate early promoter.

It is expected that the recombinant DNA molecule described above will, most typically, be incorporated into mammalian cells that constitutively express the tet repressor protein. Suitable cells may be developed by transforming a mammalian cell line, *e.g.*, U2OS cells or Vero cells, with a vector containing the tet repressor protein gene operably linked to a promoter active in the cells (*e.g.*, a CMV promoter, HSV-1 promoter or SV40 promoter). Alternatively, the DNA molecule may contain, in addition to the elements already discussed, a second promoter, preferably constitutive, operably linked to the tet repressor gene sequence. The invention encompasses, not only the DNA molecules, but also the host cells transformed with the DNAs and the recombinant proteins made by the cells.

The present invention is also directed to a method for recombinantly producing protein in which mammalian host cells are transformed with a vector containing a mammalian promoter sequence having a TATA element; at least one tet operator sequence positioned at least 6 nucleotides 3′ to the TATA element; and a gene lying 3′ to the operator and operably linked to the promoter. The gene 3′ to the operator may encode an antisense nucleic acid that inhibits the expression of a selected gene, a therapeutically active agent (e.g. a tumor suppressor or a transdominant negative mutant polypeptide of a cellular protein), a protein of interest for experimental purposes or simply a protein whose isolation is desired. In all cases where the gene encodes a protein, the operator sequence will be positioned before the translation initiation codon of the gene. The transformed cells should constitutively express the repressor protein and recombinant gene expression may be induced in the cells by introducing tetracycline. Typically, the tet operator sequence will be located between 6 and 100 nucleotides (preferably between

6 and 24 nucleotides) 3′ to the last nucleotide in the TATA element. The preferred promoter is the human CMV immediate-early promoter. It has been found that this system allows for the very tight regulation of gene expression, i.e., expression is essentially completely shut off until the inducer, tetracycline, becomes available.

The method can be used to produce recombinant protein in cultured mammalian cells or in the cells of a transgenic or non-transgenic animal. When a transgenic animal is used for production, it will most typically be a mouse and it is necessary that the cells transformed with the vector described above be embryonic stem cells. The stem cells may be engineered to express the tetracycline repressor by transforming them with the repressor gene operably linked to a promoter prior to transformation with the tet operator and recombinant gene. Alternatively, the repressor gene can be incorporated into the same DNA construct as the tet operator and placed under the control of either the same promoter as the gene encoding the recombinant protein or under the control of a separate promoter. The transformed stem cells are incorporated into a blastocyst to form a chimeric embryo, which is implanted into a pseudopregnant animal. Embryos implanted in this manner are allowed to develop into viable offspring that are screened to identify heterozygous animals expressing the recombinant gene. The heterozygous animals are then bred to produce homozygous animals that make recombinant protein in response to the administration of tetracycline.

The invention encompasses the transgenic animals made using this method and any transgenic animal that has integrated into its genome recombinant DNA containing a mammalian promoter sequence having a TATA element; at least one tet operator sequence positioned at least 6 nucleotides 3′ to the TATA element; and a gene lying 3′ to the operator and operably linked to the promoter. When the gene encodes a protein, the sequence of the operator will be positioned before the translation initiation codon of the gene. Typically, the tet operator sequence will be located between 6 and 100 nucleotides (preferably between 6 and 24 nucleotides) 3′ to the last nucleotide in the TATA element. The preferred promoter is the human CMV immediate-early promoter. In addition to the transgenic animals, the invention encompasses the recombinant proteins made by these animals.

B. Engineered Viruses and Their Uses

One particularly important use of the tet operator/repressor expression system is in the making of viruses in which replication can be controlled. The essential characteristic of these viruses is that they contain within their genome at least three related elements: a recombinant promoter having a TATA element; at least one tet operator sequence positioned at least 6 nucleotides 3′ to the TATA element; and a gene operably linked to the promoter, which lies downstream from the operator and which inhibits viral replication when expressed. Typically, the tet operator sequence will be located between 6 and 100 nucleotides (preferably between 6 and 24 nucleotides) 3′ to the last nucleotide in the TATA element. The gene lying downstream of the operator may act either by encoding a protein that inhibits viral replication or by forming a transcription product that inhibits viral replication through an antisense mechanism. When the gene encodes a protein, the tet operator sequence will be positioned upstream from the translation initiation codon. The engineered virus can be made and grown in cultured cells that constitutively express the tet repressor protein. Under these conditions, the gene that inhibits viral replication will be shut off, allowing large amounts of virus to be produced. Virus may then be collected, purified, and introduced into mammalian cells either *in vitro* or *in vivo*. Since mammalian cells do not normally make the tet repressor protein, the operator sequence will be unoccupied. As a result, the gene lying 3′ to the tet operator is expressed and viral replication is prevented.

Viruses engineered in the manner discussed above have a wide range of possible applications. First, the viruses can be used as a vehicle for delivering DNA, (*e.g.*, a gene) to mammalian cells. Under these circumstances, a second recombinant promoter will typically be incorporated into the viral genome and operably linked to the gene whose expression is desired. This second promoter may or may not, be followed by one or more tet operators lying between 6 and 100 (preferably between 6 and 24) nucleotides downstream from a TATA element in the second recombinant promoter. After having delivered the DNA to the host cell, production of new virus is inhibited due to the absence of the tet repressor protein. The gene attached to the second promoter may encode an antisense nucleic acid that inhibits the expression of a selected gene within cells; a therapeutically active protein (*e.g.*, a tumor suppressor or a transdominant negative mutant polypeptide of a cellular protein); or simply a protein that will be isolated or that is of interest for experimental reasons. The invention encompasses the method of

transforming host cells by transfecting them with the virus, the transformed host cells themselves and the recombinant proteins made by the host cells.

The viruses discussed above may also be used to immunize subjects. The great advantage of vaccines containing the engineered viruses that, because the viruses will not replicate after they are injected into subjects, the risk of active viral infection due to immunization is greatly reduced. To further ensure that virus replication will not occur, additional mutations may be introduced into the viruses, e.g. a deletion mutation may be introduced into one or more essential viral genes. In general, viruses containing such additional mutations will be preferred.

The engineered viruses also have utility in the direct treatment of patients for viral infections. The first step in this method involves transforming a second virus (i.e., a virus other than the one that has infected the patient although possibly of the same strain) by incorporating into its genome: DNA comprising a mammalian promoter with a TATA element; at least one tet operator sequence positioned at least 6 nucleotides 3′ to the TATA element (typically between 6 and 100 nucleotides 3′ to the TATA element); and a gene positioned 3′ to the operator and operably linked to the promoter. This gene should be chosen so that, when expressed, it is capable of blocking the replication of both the second virus and the virus which has infected the patient. In cases where the gene encodes a protein, the sequence of the tet operator will be positioned before the translation initiation codon of the gene. The transformed second virus is grown in host cells expressing the tet repressor protein, thereby allowing large amounts of viral progeny to be produced. Virus is collected, purified and then administered to the patient. In preferred embodiments, the tet operator is located between 6 and 24 nucleotides downstream from the last nucleotide in the TATA box and the promoter used is the human CMV immediate-early promoter.

Finally, the present invention is directed to a method for delivering a nucleic acid therapeutic agent to cells. The nucleic acid therapeutic agent may comprise either an antisense fragment that inhibits the expression of a cellular protein, or a gene that encodes a protein with a therapeutic action. The virus is engineered to contain within its genome: i) a recombinant mammalian promoter with a TATA element; at least one tet operator sequence positioned at least 6 nucleotides 3′ to the TATA element and 5′ to a translation initiation codon; and iii) a

gene positioned 3′ to the operator and operably linked to the promoter. When this gene is expressed, viral replication is inhibited. Typically, the tet operator sequence will be located between 6 and 100 nucleotides (preferably between 6 and 24 nucleotides) 3′ to the last nucleotide in the TATA element. The preferred promoter is the immediate-early promoter of human CMV. In addition, the virus must contain within its genome the nucleic acid encoding the therapeutic agent operably linked to a second promoter. This second promoter may, or may not, be followed by one or more tet operators lying, typically, between 6 and 100 (preferably between 6 and 24) nucleotides downstream from a TATA element in the second promoter.

After the preparation of the viral vector for delivering therapeutic agent, the next step in the method is to grow a large amount of the virus in a host cell that expresses the tet repressor protein. The virus grown in this manner is collected, purified and then administered to the patient. Since the patient would not normally have cells synthesizing tet repressor, replication of virus will be blocked but transcription of the nucleic acid therapeutic agent will proceed.

Brief Description of the Figures

Figure 1. Diagram of the hCMV major immediate-early enhancer-promoter and strategy for creating a tetR-responsive transcription switch. Panel A shows a DNA sequence containing two tandem tet operators used for generating the tet operator-bearing hCMV major immediate-early enhancer-promoter. Panel B shows DNA sequences surrounding the TATA element and cis-acting sequences known to interact with cellular transcription factors. A Sac I restriction site used for insertion of tet operator is underlined.

Figure 2. Insertion of tet operator sequences immediately downstream of the TATA element converts the hCMV major immediate-early enhancer-promoter to a tetR-sensitive transcription switch. Vero cells were seeded at 3 x 10^5 cells per 60 mm dish and, at 24 hours after seeding, cells were transfected with 0.5 μg of pWRG1630 or pCMVtetOEGF alone or in the presence of 1 μg, 2 μg and 3 μg of the tet repressor-expressing plasmid, pcDNA3-tetR, either in the absence or presence of tetracycline at 1 μg/ml. pUC19 vector plasmid was used to balance the pCDNA3-tetR and 3.5 μg of plasmid DNA was used in the transfection assay. Extracellular medium was collected from transfected cells every 20-24 hours and fresh growth medium was added either with or without 1 μg/ml of tetracycline. hEGF in extracellular medium collected

from 0 to 20 hours (panel A), 20 to 44 hours (panel B), and 44 to 68 hours (panel C) was determined by ELISA with the use of anti-hEGF specific monoclonal and polyclonal antibodies.

Figure 3. Release of tetR-mediated repression by tetracycline. Vero cells were transfected with either pCMVtetOEGF (0.5 μg) or pCMVtetOEGF (0.5 μg) and pcDNA3-tetR (2 μg) in the absence of tetracycline. 20 hours after transfection, extracellular medium was collected and fresh growth medium was added to the transfected cells in the absence or presence of tetracycline at 0.1 μg/ml or 1 μg/ml for an additional 24 hours. hEGF expression in extracellular medium collected from 0 to 20 hours (panel A) and 20 to 44 hours (panel B) was determined by ELISA. Two independent experiments are shown for each indicated co-transfection assay.

Figure 4. Determination of the efficacy of tetR-mediated cumulative regulation of transgene expression using luciferase as a reporter. Vero cells in 60 mm dishes were transfected with 0.5 μg pCMVtetOGL2 alone or 0.5 μg of pCMVtetOGL2 together with 2 μg of pCDN3-tetR in the absence of tetracycline from 0 to 20 hours and in the absence or presence of 1 μg/ml of tetracycline from 20 to 70 hours. At 70 hours after transfection, Vero cells were lysed in 0.5 ml of 1x luciferase lysis buffer in the presence of 0.2 mM of PMSF, 100 μg/ml of TPCK and 1 mM of leupeptin for 15 minutes at room temperature. Insoluble cellular debris was removed by centrifugation in a microcentrifuge for 20 minutes at 4°C and luciferase activity was then measured as mV per 10 μg of protein.

Figure 5. *In vivo* regulation of the hCMV major immediate-early enhancer-promoter by the tet repressor. A total of 18 partial thickness wounds (15 x 15 x 15 x 1.2 mM) were created on porcine dorsal skin. Nine wounds received 0.2 μg of pCMVtetOEGF and 0.8 μg of pcDNA3 vector DNA and the others received 0.2 μg of pCMVtetOEGF and 0.8 μg of pcDNA3-tetR by particle-mediated gene transfer. After particle bombardment, each transfected wound was enclosed in a sealed vinyl adhesive chamber containing 1.2 ml of isotonic saline in the presence of 100 units/ml penicillin and 100 μg/ml streptomycin. Wound fluid was collected from the chambers at 22, 46 and 70 hours after gene transfer and stored at -70°C. Following collection of wound fluid at each indicated time point, a new chamber was applied. Each pig was given 500 mg of tetracycline by intravenous injection at 46 hours after gene transfer. hEGF expression

in wound fluid was determined by ELISA. The vertical line associated with each bar represents standard error.

Figure 6. Inhibition of HSV-1 replication by the transdominant negative form of the UL9 peptide and reversibility using the tet repressor. Vero cells were seeded at 5 x 10^5 cells per 60 mm. At 20 to 24 hours after seeding, the cells were transfected with 0.1 μg of purified infectious HSV-1 DNA either alone or in the presence of 0.1 μg of either pCMVtetOUL9-C571 or pCMVtetOUL9-n10/C535. Transfections were carried out in the presence of 1.5 μg of pCDNA3 vector DNA or the tet repressor-expressing plasmid, pCDNA3-tetR. Fourteen hours after transfection, medium was removed followed by the addition of methylcellulose to the transfected cells at 10 ml per dish. Viral plaques were visualized by staining transfected cells with neutral red at 68 to 72 hours post-transfection and plates were counted 14 hours later.

Figure 7. Reversibility of HSV-1 replication inhibition using tetracycline. Vero cells were transfected with three different sets of DNA vectors: 1) 0.2 μg of infectious HSV-1 DNA and 2.1 μg of pCDNA3; 2) 0.2 μg of infectious HSV-1 DNA, 0.1 μg of pCMVtetOUL9-C571 and 2 μg of pCDNA3; and 3) 0.2 μg of infectious HSV-1 DNA, 0.1 μg of pCMVtetOUL9-C571 and 2 μg of pCDNA3-tetR. Transfections were carried out either in the presence or absence of tetracycline at 1 μg/ml. Sixteen hours after transfection, medium was removed from cells and 5 ml of fresh medium was added to each dish either with or without tetracycline at a concentration of 5 μg/ml. At 48 hours after transfection, cells were harvested and virus yields were determined. The results of this determination are shown in the figure.

Definitions

The description that follows uses a number of terms that refer to recombinant DNA technology. In order to provide a clear and consistent understanding of the specification and claims, including the scope be given such terms, the following definitions are provided.

Viral vector: As used herein, "viral vector" and equivalent terms refer to viruses that are utilized for transferring selected DNA or RNA sequences into a host cell. The vectors maybe utilized for the purpose of transferring DNA into cells either *in vitro* or *in vivo*. Viruses that

have been commonly used for the latter purpose include the retroviruses, adenoviruses, parvoviruses and herpes viruses.

Expression vector: This and comparable terms refer to a vector which is capable of inducing the expression of DNA that has been cloned into it after transformation into a host cell. The cloned DNA is usually placed under the control of (*i.e.*, operably linked to) certain regulatory sequences such a promoters or enhancers. Promoters sequences maybe constitutive, inducible or repressible.

Substantially pure or purified: As used herein, "substantially pure" or "purified" means that the desired product is essentially free from contaminating cellular components. Containments may include, but are not limited to, proteins, carbohydrates and lipids. One method for determining the purity of a protein or nucleic acid is by electrophoresis in a matrix such as polyacrylamide or agarose. Purity is evidence by the appearance of a single band after staining.

Host: Any prokaryotic or eukaryotic cell that is the recipient of a vector is the host for that vector. The term encompasses prokaryotic or eukaryotic cells that have been engineered to incorporated a gene in their genome. Cells that can serve as hosts are well known in the art as are techniques for cellular transformation (*see e.g.*, Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor (1989)).

Promotor: A DNA sequence that initiates the transcription of a gene. Promoters are typically found 5' to the gene and located proximal to the start codon. If a promotor is of the inducible type, then the rate of transcription increases in response to an inducing agent.

Expression: Expression is the process by which a polypeptide is produced from DNA. The process involves the transcription of the gene into mRNA and the translation of this mRNA into a polypeptide. Depending on the context in which used, "expression" may refer to the production of RNA, protein or both.

Recombinant: As used herein, the term "recombinant" refers to nucleic acid that is formed by experimentally recombining nucleic acid sequences and sequence elements. A recombinant

host would be any host receiving a recombinant nucleic acid and the term "recombinant protein" refers to protein produced by such a host.

Operably linked: The term "operably linked" refers to genetic elements that are joined in such a manner that enables them to carry out their normal functions. For example, a gene is operably linked to a promotor when its transcription is under the control of the promotor and such transcription produces the protein normally encoded by the gene.

Nucleic acid therapeutic agent: This term refers to any nucleic acid sequence which directly, or indirectly, serves as a therapeutic agent. Typically, such agents will fall into two categories. The first category encompasses antisense nucleic acids that are designed to anneal to complementary sequences within the host cell, thereby inhibiting expression. Alternatively, the term may refer to nucleic acids that encode a therapeutic protein.

Gene: As used herein, "gene" refers to the nucleic acid sequence that undergoes transcription as the result of promoter activity. A gene may code for a particular protein or, alternatively, code for an RNA sequence that is of interest in itself, e.g. because it acts as an antisense inhibitor.

Mammalian promoter: The term "mammalian promoter" refers to promoters that are active in mammalian cells. Similarly, "prokaryotic promoter" refers to promoters active in prokaryotic cells.

Essential viral gene: The term "essential viral gene" is defined as a gene that is necessary for viral replication.

Essential cellular gene: This refers to a gene that is necessary for cellular survival

Detailed Description of the Invention

The present invention is based upon the concept that is it possible to regulate mammalian gene expression using the tet operator and repressor protein. Provided that the operator is positioned at least 6 nucleotides downstream from the last nucleotide of the TATA element of

the promoter controlling expression, regulation can be accomplished without the need to fuse the repressor protein to other mammalian transcription modulators.

Although not critical, a knowledge of the basic functioning of the tetracycline resistance (tet) operon in bacteria may help in understanding the way in which the invention works. In the tet operon, a tetracycline resistance gene (tetA) and gene encoding the tet repressor protein (tetR) are both under the control of the same promotor and operator elements. In the absence of tetracycline, the tet repressor protein binds to the operator DNA sequence, thereby sterically preventing the adjacent promotor from interacting with RNA polymerase. Thus, transcription of both tetA and tetR are blocked. When the level of tetracycline within the bacterium increases, the tetracycline binds to the repressor protein causing it to detach from the operator sequence. As a result, the polymerase is able to bind to the promotor sequence and both the tetA and tetR genes are transcribed.

The strong interaction between the tet repressor protein and the tet operator has provided a mechanism for targeting eukaryotic regulatory proteins to specific sites within the genome of a cell. As discussed above, previous systems have been described in which the tet operator is positioned upstream from a mammalian gene to serve as a target for fusion proteins comprised of the tet repressor and a mammalian transcription activator or repressor. The tet repressor portion of the fusion protein binds to the operator sequence, thereby positioning it upstream from the gene to be expressed. The remaining portion of the fusion protein then serves to modulate gene expression by interacting with cellular transcription factors.

The main problem with these types of systems is that pleiotropic effects are caused by the interaction of the mammalian transcription modulator with transcriptional factors at sites distinct from the operator. Previous attempts to modulate gene expression using the tet repressor protein alone, (*i.e.*, other than as a fusion protein) have been unsuccessful (*see e.g.*, Kim, *et al.*, *J. Virol. 69*:2565-2573 (1995); Deuschle, *et al.*, *Mol. Cell. Biol. 15*:1907-1914 (1995)). It has now been discovered that successful modulation of gene expression using tetR alone can be accomplished by inserting one or more tet operators approximately 10 base pairs, a full DNA helix turn, downstream of the tet operator. Using this approach, it has been possible to tightly regulate transcription controlled by the hCMV major immediate-early enhancer-

promotor, one of the most potent and promiscuous eukaryotic elements. This can be done both *in vitro* and *in vivo*.

I. The Tet Operator as a Transcriptional Switch

In its first aspect, the present invention is directed to recombinant DNA molecules containing a mammalian promoter sequence with a TATA element. A tetracycline operator sequence is positioned at least 6 nucleotides 3' to the TATA element and is followed by a DNA sequence whose transcription is controlled by the promoter. Procedures for either synthesizing or purifying promoters, operators and other DNA sequences are well known in the art and standard techniques in molecular biology can be employed for constructing DNA molecules with appropriately arranged elements (*see e.g.*, Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Press (1989)). Examples of preferred methods are provided in the "Examples" section along with the complete sequence of the tet operator.

Any type of promoter active in mammalian cells can be used in the invention including those that are inducible, repressible or constitutive. Preferred mammalian promoters include that of the mouse metallothionein I gene (Hamer, *et al. J. Mol. Appl. Gen. 1*:273-288 (1982)); the immediate-early and TK promotor of herpes virus (Yao *et al., J. Virol. 69*:6249-6258 (1995); McKnight, *Cell 31*:355-365 (1982)); the SV 40 early promotor (Benoist, *et al., Nature 290*:304-310 (1981)); and, especially, the human CMV immediate-early promotor (Boshart, *et al. Cell 41*;521-530 (1985)). Full length or minimal promoters may be used and other regulatory elements, (*see e.g.* Figure 1) may be included. As discussed in the "Examples" section, the full human CMV major immediate-early enhancer-promotor has been successfully used in the invention and it will be understood that, unless otherwise specified, reference to the "human CMV immediate-early promoter" includes both the promoter *per se*, as well as the promoter in combination with any or all of the other transcriptional regulatory elements shown in Figure 1.

The promotor is separated from the sequence undergoing transcription by one or more tet operator sequences that begin at least 6 nucleotides downstream from the TATA element. Typically, the operator will begin at a position between 6 and 100 nucleotides (and preferably

between 6 and 24 nucleotides) downstream from the TATA element. The arrangement of these elements must not substantially interfere with the ability of the promoter to direct the transcription of the downstream sequence or the translation of the gene product.

Typically, the DNA molecule described above will be incorporated into a vector (e.g.. a plasmid or virus) which contains other transcription or translational elements. If desired, large amounts of vector DNA can be generated, (e.g., but transferring the vector into bacteria that make the repressor protein). Preferably, the vector is then transferred into a mammalian host cell which has been engineered to express the tet repressor. One way to engineer mammalian cells to express the tet repressor is to operably link the repressor gene sequence to a second promoter, incorporate this into the vector containing the tet operator and then transfer the DNA into the cells. Alternatively, cells may be transformed with an expression vector containing the tet repressor sequence prior to the transfer of the construct containing the tet operator. An example of a plasmid that has been used to produce cells expressing the tet repressor is pcDNA3-tetR (see "Examples" section).

Any method for introducing expression vectors into cells maybe used with the present invention including calcium phosphate precipitation, microinjection, electroporation, liposomal transfer, viral transfer or particle mediated gene transfer. When transfers are done to host cells *in vivo*, the preferred method of transformation is by means of a viral vector. Cells that have incorporated constructs can be identified using hybridization techniques well known in the art or by using the polymerase chain reaction (PCR) to amplify specific recombinant sequences. If the recombinant DNA transferred into the cells produces a protein that can be detected, e.g., by means of an immunological or enzymatic assay, then the presence of recombinant protein can be confirmed by introducing tetracycline into cells and then performing the assays either on the medium surrounding the cells or on cellular lysates.

In the absence of tetracycline, host cells transformed with the constructs should not express substantial amounts of recombinant DNA. Expression of recombinant DNA sequences incorporated into hosts cells is induced using either tetracycline *per se* or a tetracycline analogue. The latter is defined as any compound which is related to tetracycline in the sense that it maintains the ability to bind with specificity to the tet repressor. The dissociation

constants of such analogues should be at least 1 x 10^{-6} M and preferably greater than 1 x 10^{-9} M. Examples of analogues that can be used include, but are not limited to, those discussed by Hlavka, *et al.* ("The Tetracyclines," *in Handbook of Experimental Pharmacology* 78, Blackwood, et al. (eds.), New York (1985)) and Mitschef ("The Chemistry of Tetracycline Antibiotics," *Medicinal Res.* 9, New York (1978)). Similarly, minor modifications in the sequence of the repressor or the operator will not affect the invention provided that such modifications do not substantially reduce either the affinity or specificity of the repressor/operator interaction.

II. Method for Recombinantly Producing Protein *in Vitro* and *in Vivo*

The vectors and DNA constructs discussed above can be used as part of a method for recombinantly producing protein either *in vitro* or *in vivo*. *In vitro*, mammalian host cells are preferred for the production of protein and include U2OS cells, Vero cells, NIH-3T3 cells, CHO cells, Hela cells, LM(tk-) cells, etc. Vectors suitable for use in each of these various cell types are well known in the art (*see e.g.*, Sambrook, et al., *supra*).

The DNA constructs may also be used to produce recombinant proteins *in vivo* using both transgenic and non-transgenic animals. Although production in any type of transgenic animal is compatible with the invention, it is expected that mice will be used in most cases. Typically, mouse embryonic stem (ES) cells will be transformed with the DNA constructs and then incorporated into a developing mouse embryo. Any ES cell line which has the ability to integrate into and become part of the germ line of the developing embryo may be used, e.g., the murine cell line D3 (ATCC, 12301 Parklawn Drive, Rockville, Md., catalog no. CR 1934). The cells are cultured and prepared for DNA insertion using methods well-known in the art (*See, e.g.*, Robertson, in Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, Robertson, ed., I.R.L. Press Washington, D.C. (1987); Bradley, et al., *Current Topics in Devel. Biol 20*; 357-371 (1986); and Hogan, et al., Manipulating the Mouse Embryo: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, M.Y. (1986)). Stem cells will need to be engineered to express the tet repressor protein, and, as discussed above, this can be done either by incorporating the repressor gene into the same construct containing the tet operator or by separately transforming cells with a repressor gene-containing construct.

DNA can be incorporated into cells using any method known in the art, but most typically, this transfer will be accomplished using electroporation. If the DNA construct has been inserted into a plasmid-type vector, it is preferred that the DNA be linearized prior to transfection. Linearization can be accomplished by digesting the DNA vector with a suitable restriction endonuclease selected to cut outside of the DNA sequence to be expressed. The screening of transfected stem cells can be carried out using any of a variety of methods. For example, Southern hybridizations may be carried out using labeled probes that are specific to a sequence located within the DNA transferred into cells. Alternatively, PCR amplification can be used for selected sequences.

After embryonic stem cells have been transformed and selected, the next step is to incorporate the cells into an embryo. The preferred method for accomplishing this is by microinjection of the stem cells into an embryo at the blastocyst stage of development. In mice, blastocysts at about 3.5 days of development may be obtained by perfusing the uterus of pregnant animals. Appropriate methods for carrying this out are well known in the art (*see* Bradley, in Teratocarcinomas and Embryonic Stem Cells; A Practical Approach, (1987)). Preferred blastocysts are male and have genes s for a phenotypic marker (e.g. coat color) that is different from the phenotypic marker encoded by the stem cell genes. In this way, offspring can be easily screened.

The next step in the process of producing transgenic animals involves implanting the chimeric embryo into the uterus of a pseudopregnant animal. Such animals are typically prepared by mating females with vasectomized males of the same species. The pseudopregnant stage of the female is important for successful implantation and will vary from species to species. For mice, females about two to three days pseudopregnant should typically be used.

After chimeric embryos have been implanted into pseudopregnant animals, they are allowed to develop to term and offspring are then screened. In cases where a phenotype selection strategy has been employed, initial screening may be accomplished by simple inspection of animals for mosaic coat color or for some other readily apparent phenotypic marker. In addition, or as an alternative, chromosomal DNA may be obtained from the tissue of offspring, *e.g.*, from the tail tissue of mice, and screened for the presence of recombinant DNA using

Southern blots and/or PCR amplification. Homozygous transgenic animals may then be produced by interbreeding heterozygotes and then used to provide a continual supply of animals that are capable of expressing recombinant DNA. Expression can be controlled by maintaining the animals in the absence of tetracycline until recombinant synthesis is desired. Under these conditions, the tet repressor protein will bind to the operator sequence thereby inhibiting the activity of the recombinant promoter. Tetracycline, or a tetracycline analog, once administered to animals will readily cross cell membranes and then cause the tet repressor protein to dissociate from the operator sequence. Thus, the recombinant gene downstream from the recombinant promoter will start being transcribed.

Animals made in this manner, may be used for research purposes, *e.g.*, to study the effects of various drugs or, alternatively, they may be used for the purpose of producing recombinant protein. In the latter case, it is preferred that the recombinant genes expressed in the cells be linked to a signal sequence that causes protein to be secreted into the blood of the animals. This may then be collected to serve as a source for the purification of recombinant protein.

III. Recombinantly Engineered Virus

A. The Making of Recombinant Virus, Vaccines and Anti-viral Treatment

The tet operator/repressor regulatory system described above can be used to engineer viruses in which the production of progeny is tightly regulated. This can be done by incorporating into the viral genome a construct containing a promoter (preferably the human CMV immediate-early promoter), the tet operator sequence at a position at least 6 nucleotides 3′ to the TATA element and a gene 3′ to the operator and operably linked to the promoter. This gene inhibits viral replication when expressed and may take the form of an antisense sequence that binds to RNA encoding a protein necessary for viral replication or, alternatively, the gene may encode a protein that inhibits replication. In the latter case, the tet operator sequence will be positioned before the translation initiation codon of the gene. An example of a protein that will inhibit viral replication is the transdominant negative form of the UL9 protein of HSV-1 which binds to the HSV-1 origin of replication and, when over- expressed, blocks new viruses from being formed. Similar proteins have been found to exist in many other viruses as well.

Viruses as described above can be generated in cells that constitutively produce the tet repressor protein. Under these circumstances, the repressor will bind to the tet operator sequence and inhibit the expression of the gene downstream. Thus, viral inhibitory DNA sequences can be incorporated into the viral genome and large amounts of virus can be produced. For example, the repressor might block the synthesis of the mutant form of the UL9 protein, thereby allowing the production of HSV-1. If desired, tetracycline or a tetracycline analog may be introduced into cells. The tetracycline will bind to the repressor protein and thereby cause it to dissociate from the operator sequence. Transcription of nucleic acid from the recombinant promoter would then proceed and viral replication would be inhibited.

It should be noted that the system described above can be used both *in vitro* and *in vivo*. For example, large amount of virus can be grown by infecting cultured cells that make the tet repressor protein. The viruses can then be collected, purified and administered to a subject. Once administered, the virus delivers its DNA to the cells within the subject but, because the tet repressor protein is not present, transcription of recombinant DNA within the viral genome proceeds and viral replication is inhibited. These characteristics, in themselves, make the engineered virus particularly attractive for use in immunization procedures and in the treatment of viral diseases.

B. The Use of Engineered Virus in Immunization Procedures

Most immunization procedures are carried out by exposing a subject to a particular disease-causing agent which has been modified so that it provokes an immunological response without actually causing the disease. For example, vaccines containing either dead or attenuated virus may be given to an individual to immunize them against polio. Viruses engineered using the tet operator/repressor system can be grown in large numbers in cultured cells making the tet repressor and then administered to patients as part of a vaccine. The patients thus treated would be exposed to the proteins normally present on the virus and will therefore mount an immunological response. However, because mammalian cells do not normally make the tet repressor, the virus will not be able to replicate and full-fledged exposure to the disease will be prevented.

In order to further ensure that the virus is not made, additional mutations can be introduced into the recombinant virus. For example, a deletion mutation may be introduced into an essential viral gene. The latter virus could be made and grown in cells expressing both tetR and the wild type form of the essential viral gene.

This approach to immunization could be used for virtually all infectious viruses that have been isolated and could be applied both to the immunization of people as well as animals.

C. The Use of Engineered Virus in the Treatment of Viral Diseases

Viruses engineered using the tet operator/repressor system of the present invention can be used directly in the treatment of viral infections. For example, an HSV-1 virus could be engineered in the manner described above to contain within its genome a construct made up of a strong mammalian promoter, the tet operator sequence and the gene encoding the transdominant negative mutant form of UL9. The engineered HSV-1 could be grown in large numbers in cultured cells expressing the tet repressor and then administered to patients suffering from an HSV-1 infection. The engineered virus would enter into the patient's cells and express the transdominant negative mutant UL9 protein. This would serve to inhibit not only the replication of the engineered HSV-1 but also the HSV-1 that had originally infected the patient. In effect, the engineered virus is serving as a vehicle for delivering antiviral agents *in vivo*. Because the engineered virus shares the same cellular specificity as the infecting virus, it is ideally suited for therapy.

The animal and human viruses for which engineered virus could serve as either a vaccine or therapeutic agent include, without limitation, arboviruses; avian leukosis virus; CELO virus; Chagres virus; rhinoviruses; Coxsackie virus; hemorrhagic viruses; equine encephalomyelitis virus; hepatitis viruses; herpes viruses; infectious porcine encephalomyelitis virus; influenza viruses; Newcastle disease virus; papilloma virus; parainfluenza viruses; poliomyelitis virus; respiratory syncytial virus; Rous sarcoma virus; St. Louis encephalitis virus; dengue virus; Sendai virus; and rabies virus.

D. Engineered Viruses as Vectors for the Delivery of Nucleic Acid Therapeutics

With minor modifications, the engineered viruses discussed above can be used for delivering any type of nucleic acid therapeutic agent to cells. These agents may take the form of either antisense nucleic acids that bind to complementary sequences to inhibit their expression as proteins, or as genes encoding proteins with a therapeutic action.

The nucleic acid sequence that will be used as a therapeutic agent must be operably linked to a promoter which is active in the cells in which therapy is needed. This may either be the same promoter regulating the recombinant gene controlling viral replication or, alternatively, a second distinct promoter within the viral genome. The basic procedure to be followed in treating patients is essentially the same as that discussed above in connection with the use of engineered viruses for treating viral infections. Specifically, the virus engineered to contain nucleic acid therapeutic agent will be grown in cells that produce the tet repressor protein. Viruses made in this manner are collected, purified and administered to the subject in need of treatment. The engineered viruses then infect the subject's cells and, once inside, begin expressing both the nucleic acid inhibiting viral replication and the nucleic acid serving as a therapeutic agent. Although this system is ideally suited to gene therapy, it can also be utilized as a mechanism for delivering nucleic acids to cells *in vitro*, or as a means for attempting to engineer cells *in vivo*. For example, DNA constructs designed for homologous recombination to either replace defective counterparts or prevent abnormal gene expression may be delivered in this manner.

As discussed above, additional mutations may be introduced into an essential viral gene in order to ensure that virus is not replicated.

Examples

Example 1: Conversion of human CMV Major Immediate-early Enhancer-promoter to a Regulatory Switch Using the tet Repressor

A. Materials and Methods

Reporter and tet Expression Plasmids: Plasmid pWRG1630 is a human EGF expression plasmid in which a sequence coding for mature hEGF is controlled by the hCMV major immediate-early enhancer-promoter. There are two Sac I sites in pWRG1630 and one of these

Sac I sites is located three bases downstream of the TATA element of the hCMV major immediate-early promoter. To construct pCMVtetOEGF, the oligonucleotide:

5′ -CTCCCTATCAGTGATAGAGATCTCCCTATCAGTGATAGAGAT
CGTCGACGAGCT -3′

and its complementary sequence were annealed and purified by 15% polyacrylamide gel electrophoresis as previously described (Yao, *et al.*, *J. Virol. 68*:8158-8168 (1994)). The tetracycline (tet) operator sequence is shown in bold face (Heuer, *et al. J. Mol. Biol. 202*:407-415 (1988)) and the Sal I restriction enzyme site used for cloning analysis is underlined. The purified double stranded tet operator-containing fragment was then inserted at the Sac I site of the hCMV immediate-early promoter in plasmid pWRG1630 by partial digestion of pWRG1630 with Sac I. The insertion of a tetO sequence in pWRG1630 created a unique Sal I site and insertion of tetO in the hCMV immediate-early promoter created an Eco RI-Bam HI hCMV promoter-containing fragment of 701 base pairs. Figure 1 shows a schematic diagram of the tetO-containing hCMV immediate-early promoter in plasmid pCMVtetOEGF used in the study.

pCMVGL2 and pCMVtetOGL2 are plasmids derived from a pGL2-basic vector (Promega, Madison, WI) in which the cDNA-encoding firefly luciferase is under the control of the wild-type hCMV promoter or the tetO-bearing hCMV promoter. To generate these two plasmids, the Eco RI-Bam HI hCMV promoter-containing fragment from pWRG1630 or the hCMV-tetO promoter-containing fragment from pCMVtetOEGF was inserted into the Sma I and Bgl II site of the pGL-basic vector.

The tetracycline repressor expressing plasmid, pcDNA3-tetR, was constructed by first inserting the Bgl I-Sal I-tetR containing fragment of pSG5tetR into the Xba I and Sal I site in pGEM3Z to generate pGEM3Z-tetR. The Sal I and Kpn I -tetR fragment of pGEM3Z-tetR was then cloned into the EcoR V-Kpn I site in the pcDNA3 vector.

Cell Culture and Transfection

African green monkey kidney (Vero) cells were grown and maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum. Cells were seeded at 2 to 3 x 10^5 cells per 60 mm dish. At 20 to 24 hours post-seeding, cells were transfected with either 0.5 μg of pWRG1630 or 0.5 μg of pCMVtetOEGF in the presence of 2 μg of pUC19 vector DNA or 2 μg of pcDNA3-tetR by lipofectin-mediated transfection. The transfection was carried out in serum and antibiotic free DMEM for 16-20 hours followed by removal of the transfection medium and addition of 5 ml of normal growth medium in the presence or absence of tetracycline. The preparation of lipofectin-DNA complexes was carried out according to the procedure of the manufacturer (GIBCOBRL, Life Technologies) at 10 μl of lipofectin per 2.5 μg of plasmid DNA.

For luciferase assays, Vero cells were seeded and transfected in a manner similar to that described above, with the exception of using 0.5 μg of pCMVtetOGL2 in the presence of 2 μg of pUC19 vector DNA, or 2 μg of pcDNA3-tetR. At 20 hours after transfection, the lipofectin-plasmid DNA containing medium was removed and cells were re-fed with normal growth medium in the presence or absence of 1 μg/ml of tetracycline. Cells were harvested at 70-72 hours post-transfection and cell extracts were prepared according to the protocol described by the manufactured (Promega).

Particle-Mediated Gene Transfer:

Pigs used for *in vivo* gene transfer were domestic female Yorkshire pigs, 3 to 4 months old and weighing 40-45 kg. Partial thickness wounds (15 X 15 X 1.2 mm) were made on porcine dorsal skin with a dermatome using Halothane (1-1.5%) anesthesia in a 3:5 mixture of oxygen/nitrous oxide.

Preparation of cartridges with coated DNA-gold beads for Accell (Agracetus/Geniva, Inc.) particle-mediated gene transfer and the utilization of the Accell helium gene gun were according to the protocol provided by Geniva, Inc. (8520 University Green, Middleton, WI 53562). Each partial thickness wound was provided with 0.2 μg of hEGF expressing plasmid and 0.8 μg of pcDNA3 or 0.2 μg hEGF expressing plasmid and 0.8 μg of pcDNA3-tetR. The driving pressure used was 800 pounds per square inch (psi). Following DNA transfer, the

transfected wounds were enclosed in sealed vinyl adhesive chambers containing 1.2 ml of isotonic saline in the presence of 100 units/ml penicillin and 100 µg/ml streptomycin. Wound fluid was withdrawn from the chambers at 22 hours post-gene transfer and the transfected sites were enclosed in new chambers. Following the collection of wound fluid and application of new chambers at 46 hours after gene transfer, pigs were given 500 mg of tetracycline by intravenous injection. At 24 hours after the administration of tetracycline, wound fluid was collected and stored at -70 degrees C. Levels of EGF in wound fluid was determined by ELISA with anti-HEGF specific antibody.

ELISA:

Expression of hEGF in extracellular medium and wound fluid was determined on microtiter plates (96 wells) with the use of anti-hEGF specific monoclonal antibody (MAB236, R&D systems) as the primary coating antibody at 75 ng per well and anti-hEGF specific polyclonal antibody (sc275, Santa Cruz) as secondary antibody at 100 ng per well. The HRP-conjugated goat anti-rabbit polyclonal antibody (sc-2004, Santa Cruz) was used as tertiary antibody at 3.33 ng per well. The peroxidase assay was performed according to the procedures of the TMB peroxidase EIA substrate kit (BIO-RAD) and analyzed on a Bmax Kinetic Microplate Reader (Molecular Devices Corporation, Sunnyvale, CA). The concentration of hEGF in samples was fit to a SOFTmax 4-parameter standard curve generated with the use of recombinant hEGF (234-EG, R&D systems) in a two-fold dilution ranging from a concentration of a 2 pg to 200 pg/ml in a volume of 200 µl per well.

B. Results.

In Vitro Regulation of the hCMV Major Immediate-Early Enhancer-Promoter by the Tetracycline Repressor: The hCMV major immediate-early enhancer-promoter represents one of the most potent *cis*-regulatory units for directing the expression of transgenes in mammalian cells. In addition to the TATA element, a variety of upstream *cis*-acting element have been identified (Figure 1) and, by interacting with cellular and viral transactivators, these elements ensure highly efficient transcription directed by the TATA element. Transcription initiation requires basal transcription factors to interact with the TATA element in a coordinate fashion to form a transcription pre-initiation complex. The TATA-binding protein (TBP) is the first and only basal transcription factor to interact with DNA specifically and the binding of TBP to

the TATA element signals the transcription of the promoter. The present experiments were designed to test whether the tetracycline repressor can convert the hCMV major immediate-early enhancer-promoter into a regulatory switch by interacting with two tet operators about 10 base pairs downstream of the hCMV TATA element in plasmid pWRG1630. The tet operator sequences were positioned so that the tet repressor would bind to the same side of the DNA helix as the TATA-binding protein. Based upon their close proximity, it was hypothesized that the binding of the tet repressor to the tet operator would either block the binding of the TBP to the TATA element or interfere with the assembly of the pre-initiation complex directly.

Vero cells were transfected with 0.5 μg of pWRG1630 or pCMVtetOEGF either alone or in the presence of 1 μg, 2 μg, and 3 μg of the tet repressor expressing plasmid pcDNA3-tetR in medium either with no tetracycline or with tetracycline at a concentration of 1 μg per ml. Extracellular medium was collected from transfected cells every 20-24 hours, followed by the addition of fresh growth medium either with or without 1 μg of tetracycline. The hEGF concentration in the collected extracellular medium was determined by ELISA.

The results shown in Figure 2 demonstrate that the expression of human EGF from pWRG1630 was not affected by the presence of pcDNA3-tetR and that the insertion of the tet-operator containing sequence near the hCMV major immediate-early enhancer-promoter has no effect on human EGF expression in the absence of tetR. Expression of EGF from pCMVtetOEGF was significantly reduced in the presence of tetR in a dose and time dependent manner. In the presence of 3 μg of tetR repressor-expressing plasmid, i.e,. pcDNA3-tetR, EGF expression from pCMVtetOEGF was repressed approximately 200 fold at 20 hours post-transfection, 1000 fold at 20-24 hours post-transfection, and 3500 at 44-68 hours post-transfection in the absence of tetracycline. Little or no repression was observed in the presence of tetracycline. In the presence of 2 μg of pcDNA3-tetR, approximately a 100-fold, 600-fold, and 2000-fold inhibition of expression was detected at 0-20 hours, 20-44 hours and 44-48 hours post-transfection. In the presence of 1 μg of pcDNA3-tetR, approximately a 60-fold, 100-fold, and 200-fold reduction in the synthesis of human EGF was observed at the three indicated time points. There was no human EGF expression in mock transfected Vero cells.

To test if tetR-mediated repression can be efficiently reversed by tetracycline, Vero cells were transfected with either pCMVtetOEGF alone or pCMVtetOEGF and pcDNA3-tetR in the

absence of tetracycline from 0-20 hours (Figure 3A) and in the presence or absence of tetracycline from 20-44 hours (Figure 3B). The results demonstrate that the repression observed from 0-20 hours post-transfection can be efficiently reversed by the presence of 1 μg/ml of tetracycline while 0.1 μg/ml of tetracycline is not sufficient to reverse tetR-mediated repression under the conditions tested. Consistent with the experiments presented in Figure 2, the data demonstrated that the basal promoter activity of pCMVtetOEGF was reduced 100-200 fold during 0-20 hours post-transfection, and about 500 fold from 20-44 hours post-transfection in the presence of 2 μg of pcDNA3-tetR.

Having demonstrated the kinetics of tetR-mediated regulation of the hCMV major immediate-early enhancer-promoter with a secretable peptide, human EGF, the ability of this system to regulate the expression of a non-secretable polypeptide, firefly luciferase was tested. Figure 4 shows the results of two independent experiments in which Vero cells were either transfected with 0.5 μg of pCMVtetOGL2 alone, or co-transfected with 0.5 μg of pCMVtetOGL2 and 2 μg of pcDNA3-tetR in the presence or absence of 1 μg/ml of tetracycline. The levels of luciferase expression from pCMVtetOGL2 were decreased at least 100-fold in the presence of the tetR-expressing plasmid, pcDNA3-tetR, and it was found that this repression could be released efficiently by tetracycline. The level of luciferase expression from the wild-type hCMV immediate-early enhancer-promoter are not affected by the presence of pcDNA3-tetR. When similar experiments were performed on HeLa cells, a 40-50-fold repression was detected at 68-72 hours post-transfection. This indicates that, like the tetR-VP16 based activating system, the efficiency of the tetR-based repression is cell type dependent.

In Vivo Regulation of the hCMV Major Immediate-Early Enhancer-Promoter by the Tetracycline Repressor.: The data presented above demonstrate that: (1) the tet repressor is capable of acting as a potent sequence-specific trans-repressor in cultured mammalian cells; and (2) that the insertion of two tandem operators about 10 base pairs downstream of the promoters TATA element, converts the promoter into an effective tetracycline-dependent transcriptional switch. To test if this tetR-tet operator regulatory unit is functional *in vivo*, partial thickness wounds were created on porcine dorsal skin, with nine wounds receiving 0.2 μg of pCMVtetOEGF and 0.8 μg of pcDNA3 vector DNA and nine wounds receiving 0.2 μg of pCMVtetOEGF and 0.8 μg of pcDNA3-tetR per wound. As shown in Figure 5, human EGF

expression in partial thickness wounds co-transfected with pcDNA3-tetR was significantly lower than that observed in partial thickness wounds co-transfected with pcDNA3 vector plasmid in the absence of tetracycline. A 13-fold repression was detected one day after gene transfer.

Notably, although human EGF expression in pCMVtetOEGF transfected wounds was increased approximately 3-fold from day one to day two post-gene transfer, yields of human EGF in partial thickness wounds co-transfected with pcDNA3-tetR were reduced 1.5-fold. Collectively, in the presence of tetR, levels of human EGF expression were repressed approximately 55-fold at day two post-gene transfer in the absence of tetracycline. It is of particular significance that, upon receiving tetracycline through intravenous injection from day 2 to day 3 post-gene transfer, the tetR-mediated repression was released as evidenced by a 4-fold increase of human EGF expression in wounds receiving both pCMVtetOEGF and pcDNA3-tetR. In wounds transfected with pcMVtetOEGF alone, there was about a 4-fold reduction in EGF expression from day 2 to day 3 post-gene transfer. This observation proves the feasibility of using this regulatory switch in controlling the expression of transgenes in gene therapy.

C. Discussion

Regulation of transgene expression in target cells represents one of the most critical and challenging aspects of gene therapy. Using the hCMV major immediate-early enhancer-promoter as a prototype mammalian cell promoter, it has been demonstrated that, placing tetracycline operators 10 base pairs of the TATA element, enables the tetracycline repressor to function as a potent repressor of gene expression in mammalian cells.

Recently, by fusing the KRAB repressor domain of the human KOX1 zinc-finger protein with the tet repressor and inserting DNA sequences encoding seven tet operators 685 base pairs upstream of the transcription initiation site, it has been shown that the tet-KRAB chimeric protein, but not tetR alone, can suppress the hCMV major immediate-early enhancer-promoter approximately 10-15-fold in HeLa cells in a transient expression assay using luciferase as a reporter (Deuschle, *et al.*, *Mol. Cell. Biol. 15*:1907-1914 (1995)). Using a different strategy, in which tet operators were inserted 10 base pairs, a full helix turn, downstream of the TATA

element, it has been shown that the hCMV major immediate-early enhancer-promoter can be tightly regulated by tetR alone. Based on the study of Heuer & Hillen (*J. Mol. Biol. 202*:407-415 (1988)) it was hypothesized that this specific design would place the tet repressor on the same side of the DNA helix as TBP and the binding of tetR to the tet operator provides a direct steric block for TBP. Using hEGF as a secretable promoter, the kinetics of tetR-mediated repression was explored. Close to a 4000-fold repression was observed *in vitro* at 3 days post-transfection. Combining a porcine wound model with particle-mediated gene transfer, this study has provided a direct *in vivo* confirmation of this tetR-mediated regulatory switch in fine tuning the expression of transgenes for gene therapy.

Unlike other tet repressor/operator regulatory systems, e.g., the tetR-VP16 based activating and tetR-KRAB repressor system, the regulatory switch disclosed herein does not require the use of tetracycline repressor/mammalian cell transactivator or repressor fusion proteins to achieve its effects. Thus, the potential pleiotropic effects on the expression of cellular genes caused by cellular transcription factors are minimal and higher levels of expression of the regulator (i.e., the tetracycline repressor) can be achieved. Notably, the efficacy of the tetR-mediated regulatory switch was found to vary significantly *in vitro* as compared to *in vivo*. This apparent difference can probably be explained by: 1) differences in the means of gene transfer which may lead to different co-transfection efficiency; and 2) differences in cell types.

Example 2: Viral Replication Switch

To test whether the tetR-regulated transcription switch discussed above can be converted into a novel viral replication switch to regulate de novo viral production in a reversible fashion and produce a transdestructive recombinant virus, the following experiments were performed using herpes simplex virus type 1 as a prototype.

A. Construction of Trans-dominant Negative HSV-1 UL9 Mutant Polypeptide Expressing Plasmids

The UL9 protein is one of the seven HSV-1 essential gene products that are directly involved in viral replication. UL9 binds specifically to the HSV-1 origin of DNA replication. It is a nuclear phosphoprotein 851 amino acids in length. Studies have shown that the C-

terminal amino acids 535-851 of UL9 contain the DNA binding domain of the protein and, when over-expressed, it can block virus DNA replication in a dominant negative fashion.

In order to clone, the C-terminal 317 amino acids of UL9 and place it under the control of the tet operator-containing hCMV major immediate-early enhancer-promotor, the Bam HI - Not I EGF-containing fragment in plasmid pCMVtetOEGF was replaced by the Bam HI - EcoR V UL9-containing fragment from plasmid pSP6UL9. The resulting plasmid was designated pCMVtetOUL9-C571 and expresses the C-terminal amino acids 571-851 of UL9.

To construct plasmid pCMVtetOUL9-n10/C535, a plasmid expressing a UL9 protein fragment containing amino acids 1 to 10 of UL9 and amino acids 535 to 851 of UL9, a double stranded oligo encoding the first 10 amino acids of the UL9 protein followed by amino acids Thr-Met-Gly was inserted into the Bam HI site of pCMVtetOUL9-C571. Plasmid pCMVtetOUL9-C535C, which expresses the C-terminal amino acids 535 to 851 of UL9, was constructed by the religation of Bam HI- Kpn I digested pCMVtetOUL9-n10/C535.

B. Transient Inhibition Analysis of HSV-1 Replication

To test if the mutant UL9 polypeptides encoded by pCMVtetOUL9-C571 and pCMVtetOUL9-n10/C535 can function as trans-dominant negative mutant polypeptides inhibiting HSV-1 replication, and, most importantly, to test whether an inhibitory effect can be regulated by the tet repressor, Vero cells were seeded at 5×10^5 cells per 60 nm. At 20 to 24 hours after seeding, the cells were transfected with 0.1 micrograms of purified infectious HSV-1 DNA alone or co-transfected with 0.1 micrograms of pCMVtetOUL9-C571 or pCMVtetOUL9-n10/C535 in the presence of 1.5 micrograms of pcDNA3 vector DNA or the tet repressor- expressing plasmid, pcDNA3-tetR, by lipofectin. At 14 hours post transfection, the lipofectin-DNA containing transfection medium was removed followed by addition of methylcellulose to the transfected cells at 10 ml per dish. Viral plaques were visualized by staining transfected dishes with neutral red at 68 to 72 hours post transfection and counting 14 hours later. As shown in Figure 1, co-transfection of infectious HSV-1 DNA with pCMVtetOUL9-C571 reduces the viral placque forming efficiency approximately 30 fold. When co-transfected with pCMVtetOUL9-n10/C535C, the placque forming efficiency of infectious HSV-1 DNA was reduced at least 100 fold. Significantly, both C571- and

n10/C535C- mediated repression of HSV-1 DNA replication can be efficiently silenced by the tet repressor. When a similar experiment was performed with pCMVtetOUL9-C535C, the placque formation of HSV-1 DNA was reduced at least 200 fold and again, this C535C-mediated repression can be efficiently reversed by tetR.

Having demonstrated that the inhibitory effects of the trans-dominant negative C-terminal UL9 polypeptides on HSV-1 replication can be efficiently silenced by the tet repressor, the specificity of this tetR related viral replication switch was further investigated. Vero cells were transfected with: 1) 0.2 micrograms of infectious HSV-1 DNA and 2.1 micrograms of pcDNA3; 2) 0.2 micrograms of infectious HSV-1 DNA, 0.1 micrograms of pCMVtetOUL9-571 and 2 micrograms of pcDNA3; and 3) 0.2 micrograms of infectious HSV-1 DNA, 0.1 micrograms of pCMVtetOUL9-C571 and 2 micrograms of pcDNA3-tetR. Transfections were carried out either in the absence or the presence of tetracycline at 1 microgram per ml. At 16 hours post transfection, the transfection medium was removed and 5 ml of fresh medium was added to each dish with either no tetracycline or tetracycline at a concentration of 5 micrograms per ml. At 48 hours post transfection, cells were harvested and virus yields were determined. The data presented in Figure 2 demonstrate that: 1) C571-mediated expression of HSV-1 replication can be reversed by the tet repressor; and 2) this tetR-regulated reversion of HSV-1 replication is tetracycline specific as evidenced by the effect of pCMVtetOUL9-571 on HSV-1 placque forming units was significantly reduced in the presence of tetracycline.

Collectively, these observations demonstrate that, by combining transdominant negative mutant viral polypeptides with the tetR-regulated potent mammalian transcription switch, a novel viral replication switch can be generated. In principle, any polypeptide or antisense RNA that is capable of inhibiting viral productive infection can be incorporated into this novel viral replication switch. Using this switch, a trans-destructive or inhibitory viral vector can be generated while tetR is not present in the viral genome. This trans-inhibitory viral vector is not only capable of serving as a vehicle for *in vivo* gene transfer, but is also capable of inhibiting the endogenous and/or latent virus replication. This invention can also be used for generating a viral vaccine which not only is capable of inducing an effective host immune response, but which is also able to function as a therapeutic agent helping to eliminate endogenous viral infection when encountered within the same cell.

All references cited herein are fully incorporated by reference. Having now fully described the invention, it will be understood by those of skill in the art that the invention may be practiced and wide and equivalent range of conditions, parameters and the like, without affecting the spirit or scope of the invention or any embodiments thereof.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Brigham and Women's Hospital

(ii) TITLE OF INVENTION: Tetracycline Repressor Regulated Mammalian Cell Transcription Switch and Viral Replication Switch

(iii) NUMBER OF SEQUENCES: 1

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Vinson & Elkins L.L.P.

(B) STREET: 1455 Pennsylvania Avenue, N.W.

(C) CITY: Washington

(D) STATE: D.C.

(E) COUNTRY: U.S.

(F) ZIP: 20004-1008

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:

(B) FILING DATE:

(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Sanzo, Michael A.

(B) REGISTRATION NUMBER: 36,912

(C) REFERENCE/DOCKET NUMBER: BRI331/87001

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (202) 639-6585

(B) TELEFAX: (202) 639-6604

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 54 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTCCCTATCA GTGATAGAGA TCTCCCTATC AGTGATAGAG ATCGTCGACG AGCT 54

What is Claimed is:

1. A recombinant DNA molecule comprising:
   a) a mammalian promoter sequence having a TATA element;
   b) at least one tet operator sequence positioned at least 6 nucleotides 3′ to the TATA element; and
   c) a gene lying 3′ to said operator and operably linked to said promoter.

2. The recombinant DNA of claim 1, wherein said tetracycline operator sequence is positioned between 6 and 24 nucleotides 3′ to said TATA element.

3. The DNA molecule of claim 1, wherein said promoter is the human cytomegalovirus (hCMV) immediate-early promoter.

4. The DNA molecule of claim 1, further comprising a gene encoding the tet repressor protein.

5. A host cell transformed with a vector comprising the DNA molecule of claim 1.

6. Recombinant protein made by the host cell of claim 5.

7. A method for recombinantly producing protein in a mammalian cell that makes the tet repressor protein, said method comprising:
   a) transforming said mammalian cell with a vector comprising:
      i) a mammalian promoter sequence having a TATA element;
      ii) at least one tet operator sequence positioned at least 6 nucleotides 3′ to the TATA element; and
      iii) a gene lying 3′ to said tet operator and operably linked to said promoter;
   b) introducing tetracycline into the transformed cells of step a) to induce the expression of said gene.

8. The method of claim 7, wherein said tet operator sequence is positioned between 6 and 24 nucleotides 3′ to said TATA element.

9. The method of claim 7, wherein said promoter is the human CMV immediate-early promoter.

10. The method of claim 7, wherein said mammalian cell is an embryonic stem cell and, prior to the introduction of tetracycline to induce gene expression, the method further comprises:
    (i) incorporating said stem cell into a blastocyst to form a chimeric embryo;
    (ii) implanting said chimeric embryo into a pseudopregnant animal;
    (iii) allowing said chimeric embryo to develop into a viable offspring;
    (iv) screening offspring to identify heterozygous animals expressing said gene; and
    (v) breeding said heterozygous animals to produce homozygous transgenic animals producing said protein.

11. A transgenic animal made by the method of claim 10.

12. A transgenic animal wherein said animal has integrated into its genome recombinant DNA comprising:
    a) a mammalian promoter sequence having a TATA element;
    b) at least one tet operator sequence positioned at least 6 nucleotides 3′ to the TATA element; and
    c) a gene lying 3′ to said operator and operably linked to said promoter.

13. The recombinant DNA molecule of claim 12, wherein said promoter is the human cytomegalovirus (hCMV) immediate-early promoter.

14. The recombinant DNA of claim 12, further comprising a gene encoding the tet repressor protein.

15. Recombinant protein made by the transgenic animal of claim 12.

16. A recombinantly engineered virus comprising within its genome:
    a) a recombinant promoter having a TATA element;
    b) at least one tet operator sequence positioned at least 6 nucleotides 3′ to the TATA element; and
    c) a gene lying 3′ to said operator and operably linked to said promoter, wherein said gene inhibits the replication of said virus when expressed.

17. The virus of claim 16, further comprising one or more mutations in at least one essential viral gene.

18. The virus of claim 16, wherein said tet operator element is positioned between 6 and 24 nucleotides 3′ to said TATA element.

19. The virus of claim 16, wherein said promoter is the human CMV immediate-early promoter.

20. The virus of claim 16, further comprising:
    a) a second recombinant promoter located within the viral genome; and
    b) a second recombinant gene operably linked to said second recombinant promoter.

21. The virus of claim 20, further comprising one or more mutations in an essential viral gene.

22. The virus of claim 20, further comprising at least one tet operator sequence lying at least 6 nucleotides 3′ to a TATA element in said second recombinant promoter and 5′ to said second recombinant gene.

23. A host cell made by transfecting a cell with the virus of claim 16.

24. Recombinant protein made by the host cell of claim 23

25. A vaccine comprising the virus of claim 16.

26. A method for treating a patient for an infection by a first virus, comprising:
    a) transforming a second virus by incorporating into its genome DNA comprising:
        i) a mammalian promoter having a TATA element;
        ii) at least one tet operator sequence positioned at least 6 nucleotides 3′ to the TATA element; and
        iii) a gene positioned 3′ to said operator and operably linked to said promoter, wherein said gene, when expressed, is capable of blocking the expression of both said first virus and said second virus;
    b) growing the transformed second virus of step a) in a host expressing the tet repressor protein;
    c) collecting and purifying the virus grown in step b); and
    d) administering the virus collected and purified in step c) to said patient.

27. The method of claim 26, wherein said operator is between 6 and 24 nucleotides 3′ to said TATA element.

28. The method of claim 26, wherein said promoter is the human CMV immediate-early promoter.

29. The method of claim 26, wherein step a) further comprises:
    iv) introducing one or more mutations in an essential viral gene.

30. A method for delivering a nucleic acid therapeutic agent to cells, comprising:
    a) preparing a virus to serve as a vector, wherein said virus is engineered to contain within its genome:
        i) a recombinant mammalian promoter having a TATA element;
        ii) at least one tet operator sequence positioned at least 6 nucleotides 3′ to the TATA element; and
        iii) a gene positioned 3′ to said operator and operably linked to said promoter, wherein said gene encodes a protein capable of inhibiting the replication of said virus;
        iv) said nucleic acid therapeutic agent, operably linked to a second promoter;

b) growing the virus prepared in step a) in host cells expressing the tet repressor protein;

c) collecting and purifying the virus grown in step b); and

d) administering the virus collected and purified in step c) to said patient.

31. The method of claim 30, wherein said virus further comprises at least one tet operator sequence lying at least 6 nucleotides 3′ to a TATA element in said second recombinant promoter and 5′ to said second recombinant gene.

32. The method of claim 30, wherein said tet operator sequence is positioned between 6 and 24 nucleotides 3′ to said TATA element.

33. The method of claim 30, wherein said recombinant mammalian promoter is the human CMV immediate-early promoter.

34. The method of claim 30, wherein said nucleic acid therapeutic agent acts as an antisense inhibitor of gene expression.

35. The method of claim 30, wherein said nucleic acid therapeutic agent encodes a protein with a therapeutic action.

36. The method of claim 30, wherein step a) further comprises:

iv) introducing one or more mutations in an essential viral gene.

FIG. 1A


tet-operator
CTCCCTATCAGTGATAGAGATCTCCCTATCAGTGATAGAGATCGTCGACGAGCT
TCGAGAGGGATAGTCACTATCTCTAGAGGGATAGTCACTATCTCTAGCAGCTGC
TATATAAGCAGAGCTC

FIG. 1B

HCMV major IE promoter/enhancer
-500
-400
-300
-200
-100
+1
Cis-acting element
Sp 1 binding site
CAAT box
16 bp repeat
18 bp repeat
19 bp repeat
21 bp repeat

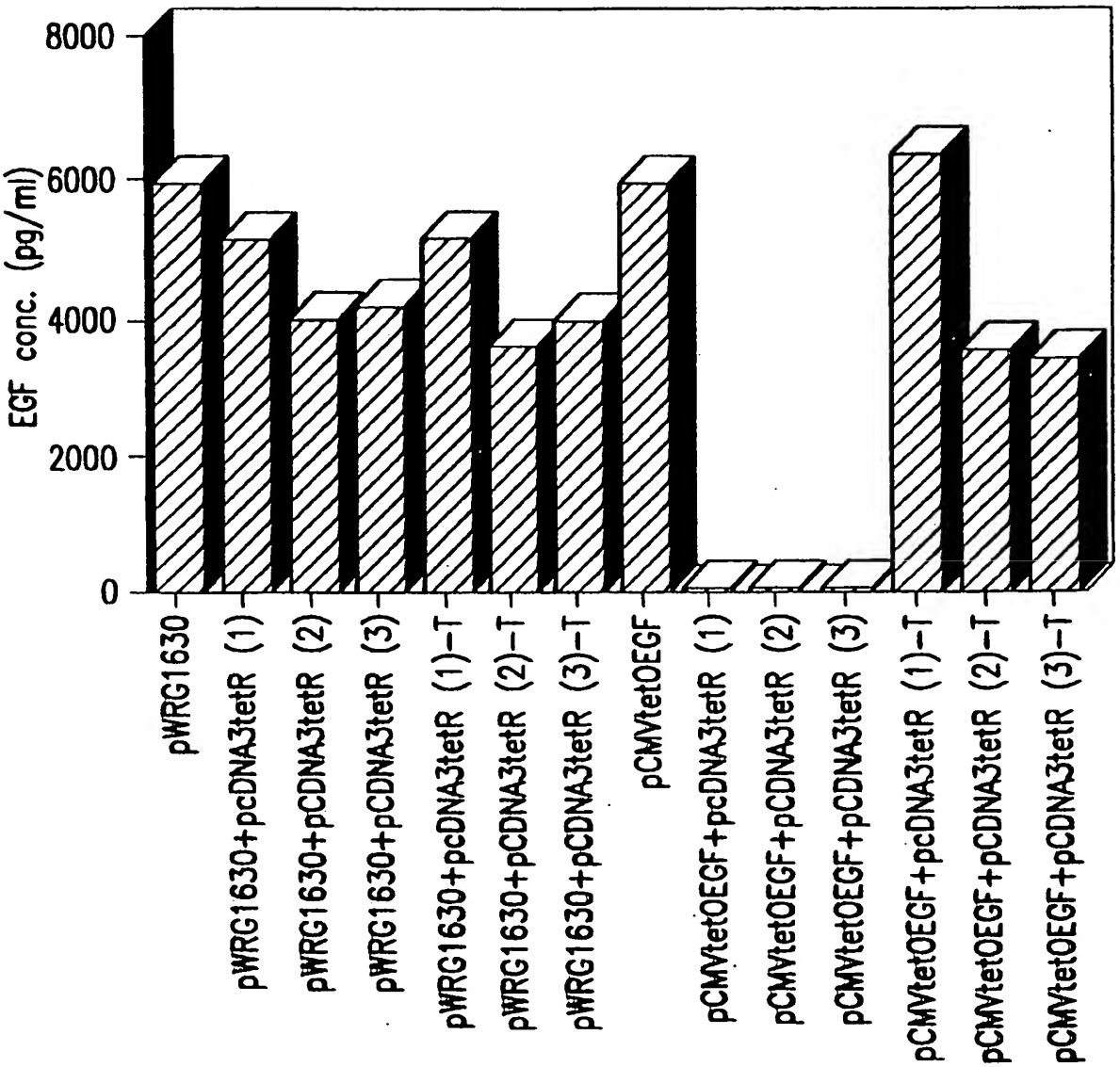

8000
6000
4000
2000
0
EGF conc. (pg/ml)
pWRG1630
pWRG1630+pcDNA3tetR (1)
pWRG1630+pcDNA3tetR (2)
pWRG1630+pcDNA3tetR (3)
pWRG1630+pcDNA3tetR (1)-T
pWRG1630+pcDNA3tetR (2)-T
pWRG1630+pcDNA3tetR (3)-T
pCMVtetOEGF
pCMVtetOEGF+pcDNA3tetR (1)
pCMVtetOEGF+pcDNA3tetR (2)
pCMVtetOEGF+pcDNA3tetR (3)
pCMVtetOEGF+pcDNA3tetR (1)-T
pCMVtetOEGF+pcDNA3tetR (2)-T
pCMVtetOEGF+pcDNA3tetR (3)-T

FIG. 2A

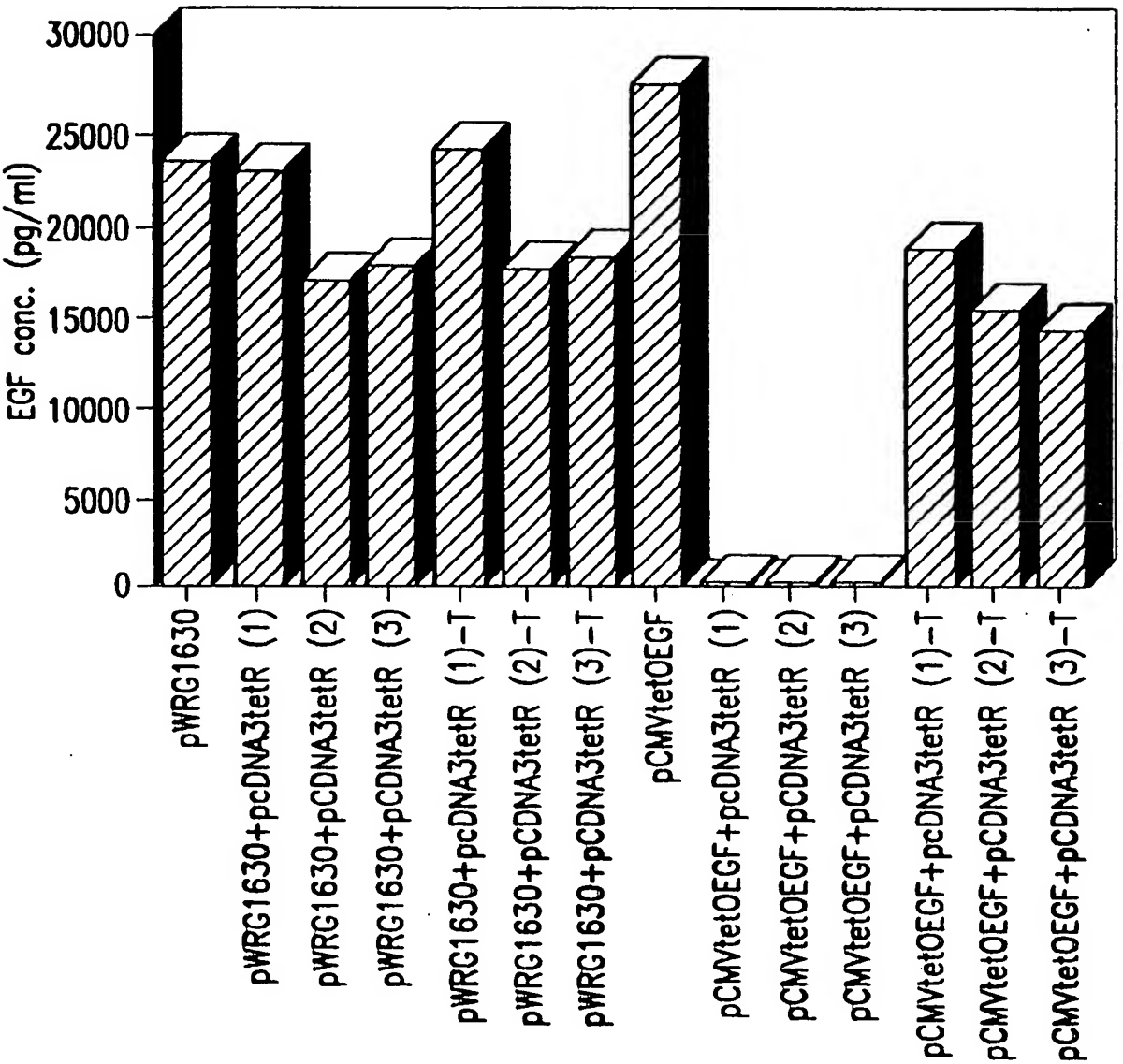

EGF conc. (pg/ml)
30000
25000
20000
15000
10000
5000
0
pWRG1630
pWRG1630+pcDNA3tetR (1)
pWRG1630+pcDNA3tetR (2)
pWRG1630+pcDNA3tetR (3)
pWRG1630+pcDNA3tetR (1)-T
pWRG1630+pcDNA3tetR (2)-T
pWRG1630+pcDNA3tetR (3)-T
pCMVtetOEGF
pCMVtetOEGF+pcDNA3tetR (1)
pCMVtetOEGF+pcDNA3tetR (2)
pCMVtetOEGF+pcDNA3tetR (3)
pCMVtetOEGF+pcDNA3tetR (1)-T
pCMVtetOEGF+pcDNA3tetR (2)-T
pCMVtetOEGF+pcDNA3tetR (3)-T

FIG. 2B

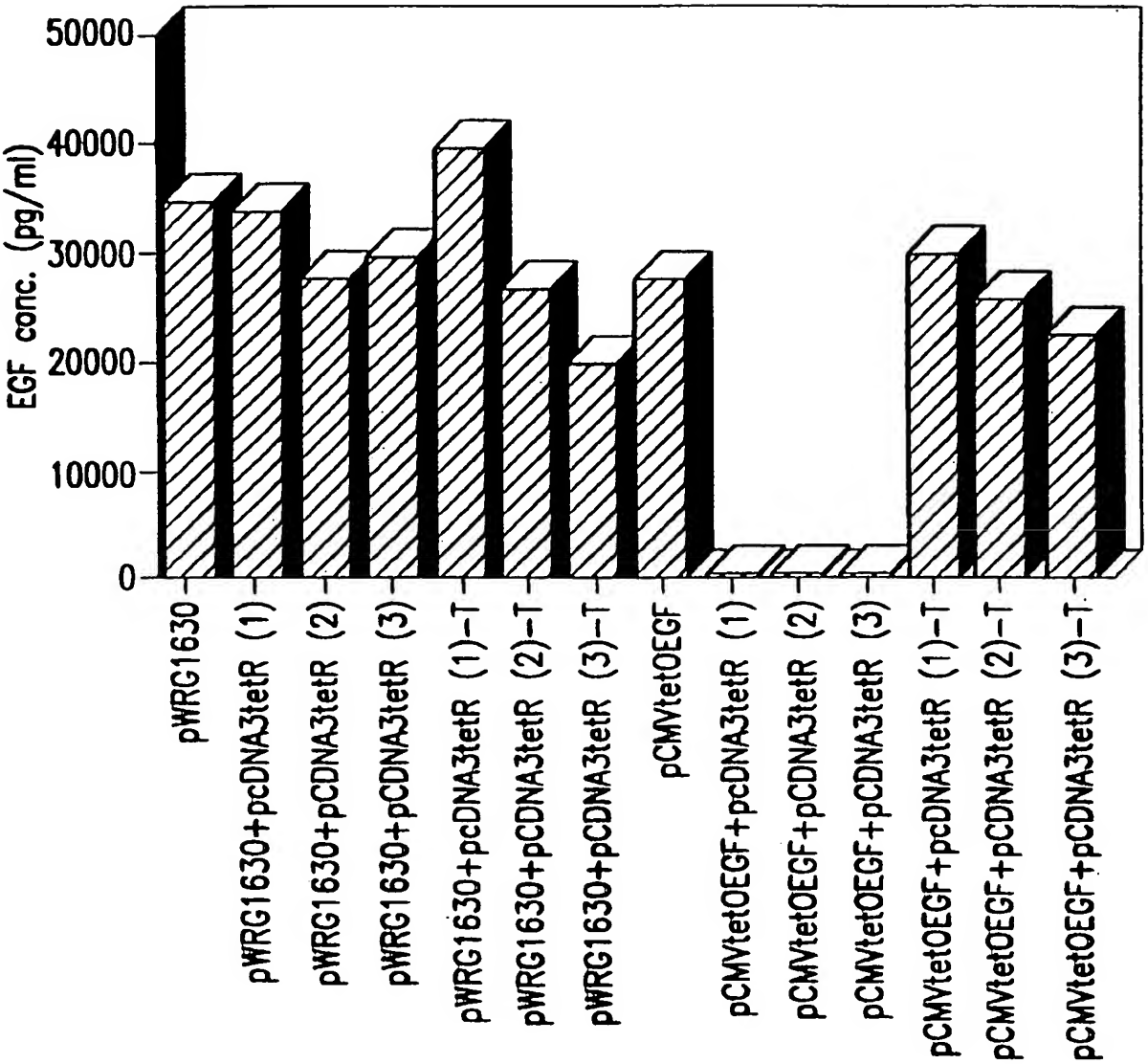

EGF conc. (pg/ml)
50000
40000
30000
20000
10000
0
pWRG1630
pWRG1630+pcDNA3tetR (1)
pWRG1630+pcDNA3tetR (2)
pWRG1630+pcDNA3tetR (3)
pWRG1630+pcDNA3tetR (1)-T
pWRG1630+pcDNA3tetR (2)-T
pWRG1630+pcDNA3tetR (3)-T
pCMVtetOEGF
pCMVtetOEGF+pcDNA3tetR (1)
pCMVtetOEGF+pcDNA3tetR (2)
pCMVtetOEGF+pcDNA3tetR (3)
pCMVtetOEGF+pcDNA3tetR (1)-T
pCMVtetOEGF+pcDNA3tetR (2)-T
pCMVtetOEGF+pcDNA3tetR (3)-T

FIG. 2C

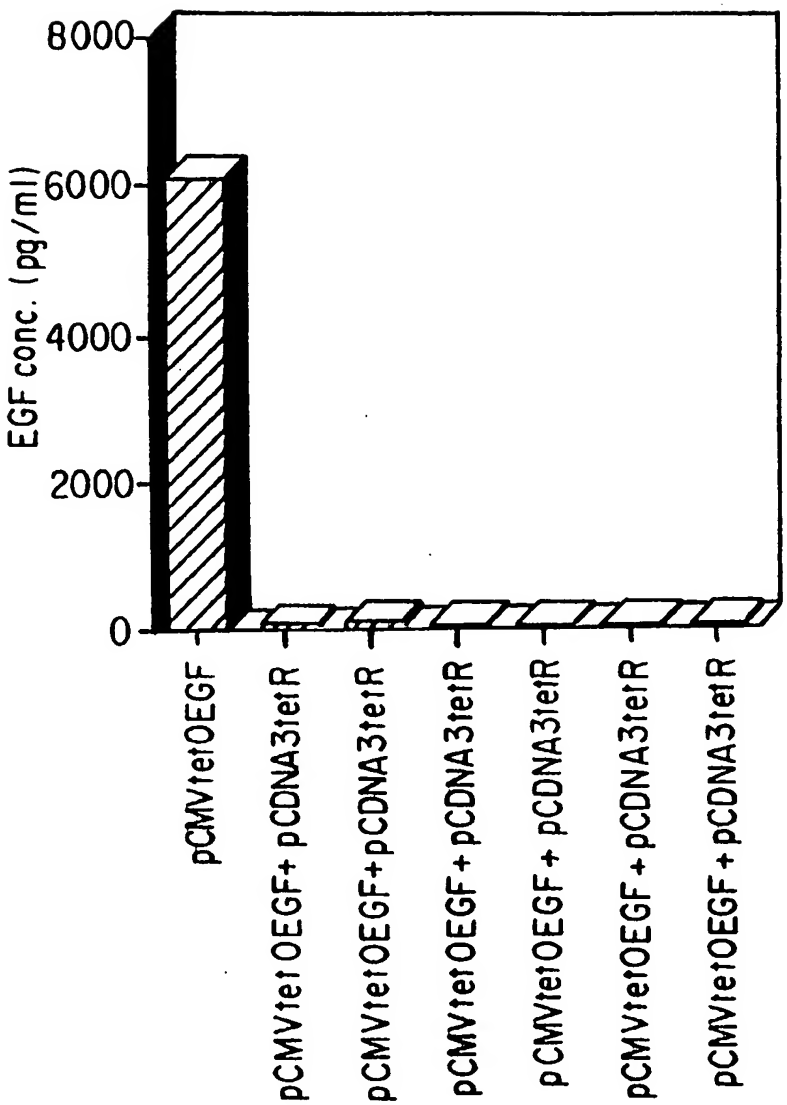

8000
6000
4000
2000
0
EGF conc. (pg/ml)
pCMVtetOEGF
pCMVtetOEGF+pCDNA3tetR
pCMVtetOEGF+pCDNA3tetR
pCMVtetOEGF+pCDNA3tetR
pCMVtetOEGF+pCDNA3tetR
pCMVtetOEGF+pCDNA3tetR
pCMVtetOEGF+pCDNA3tetR

FIG. 3A

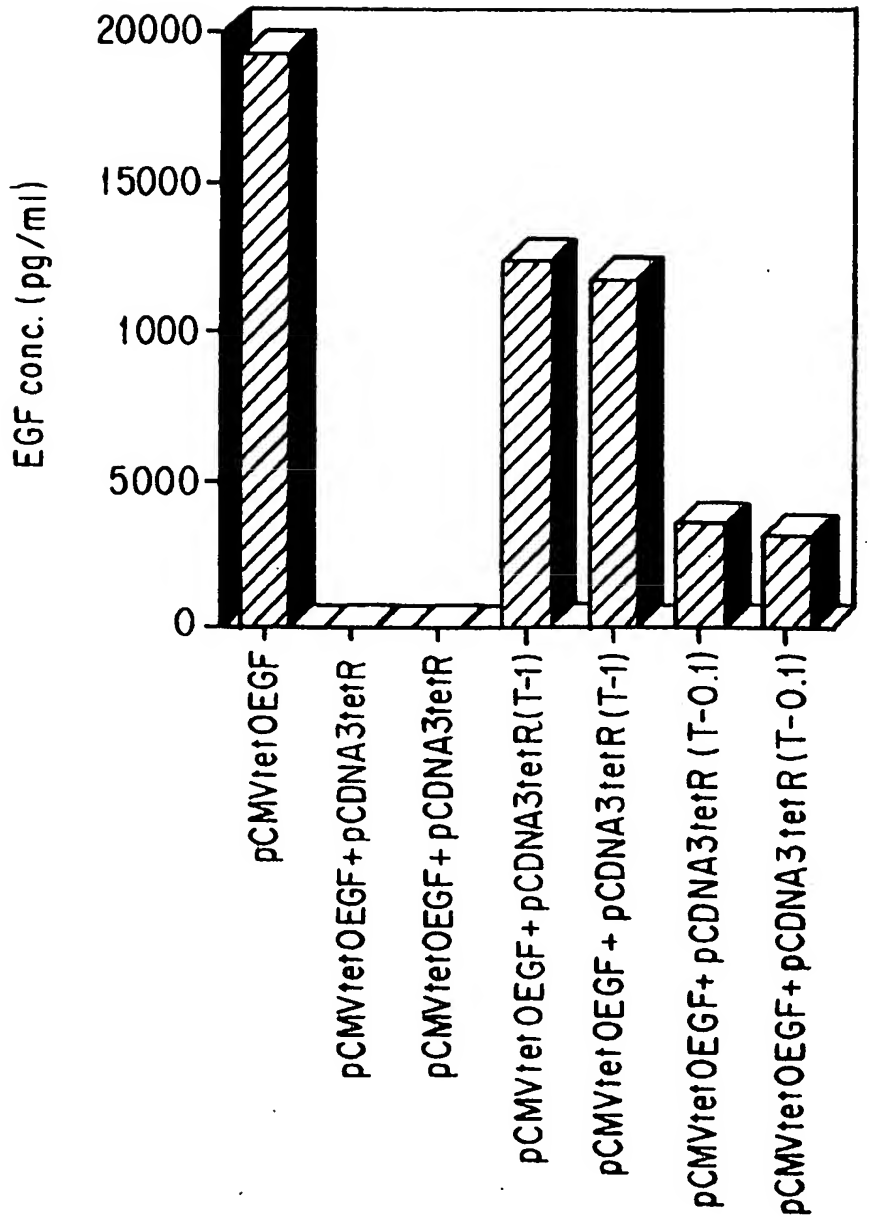

20000
15000
1000
5000
0
EGF conc. (pg/ml)
pCMVtetOEGF
pCMVtetOEGF+pCDNA3tetR
pCMVtetOEGF+pCDNA3tetR
pCMVtetOEGF+pCDNA3tetR(T-1)
pCMVtetOEGF+pCDNA3tetR(T-1)
pCMVtetOEGF+pCDNA3tetR(T-0.1)
pCMVtetOEGF+pCDNA3tetR(T-0.1)

FIG. 3B

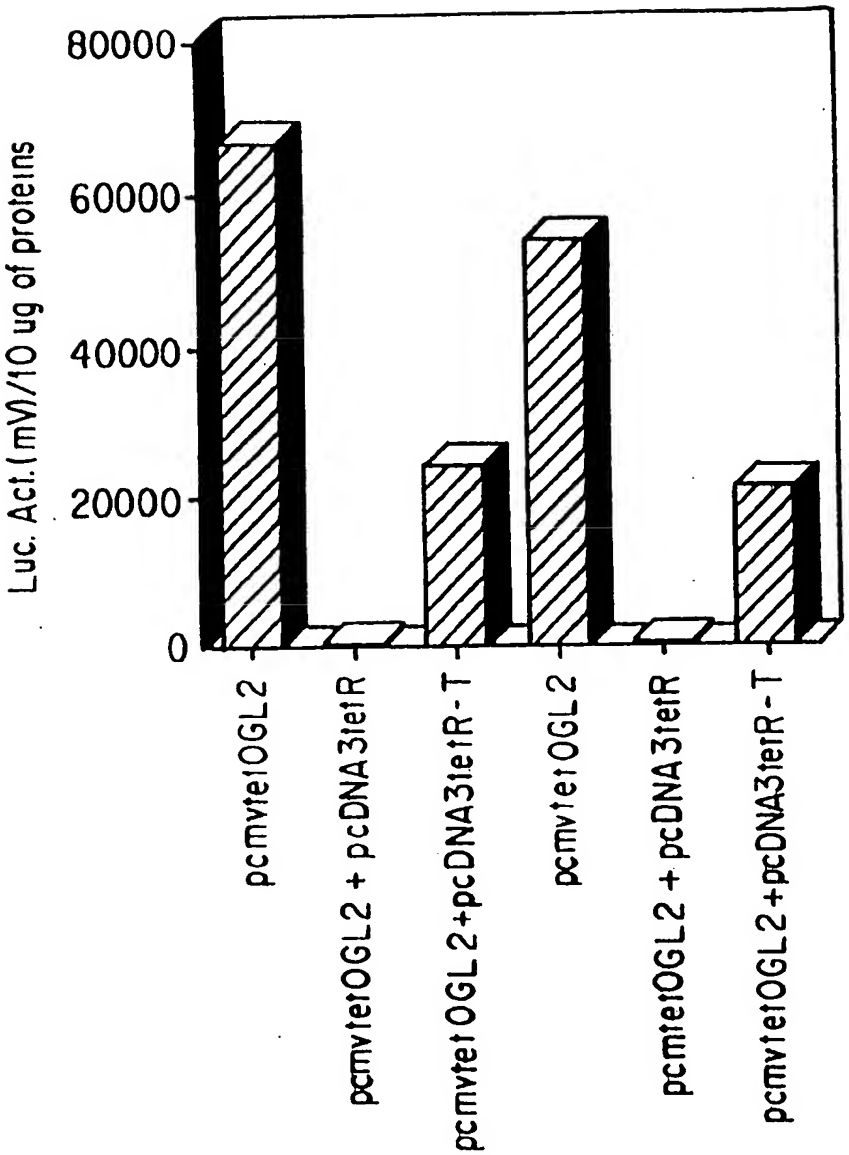

80000
60000
40000
20000
0
Luc. Act.(mV)/10 ug of proteins
pcmvtetOGL2
pcmvtetOGL2 + pcDNA3tetR
pcmvtetOGL2+pcDNA3tetR-T
pcmvtetOGL2
pcmtetOGL2 + pcDNA3tetR
pcmvtetOGL2+pcDNA3tetR-T

FIG. 4

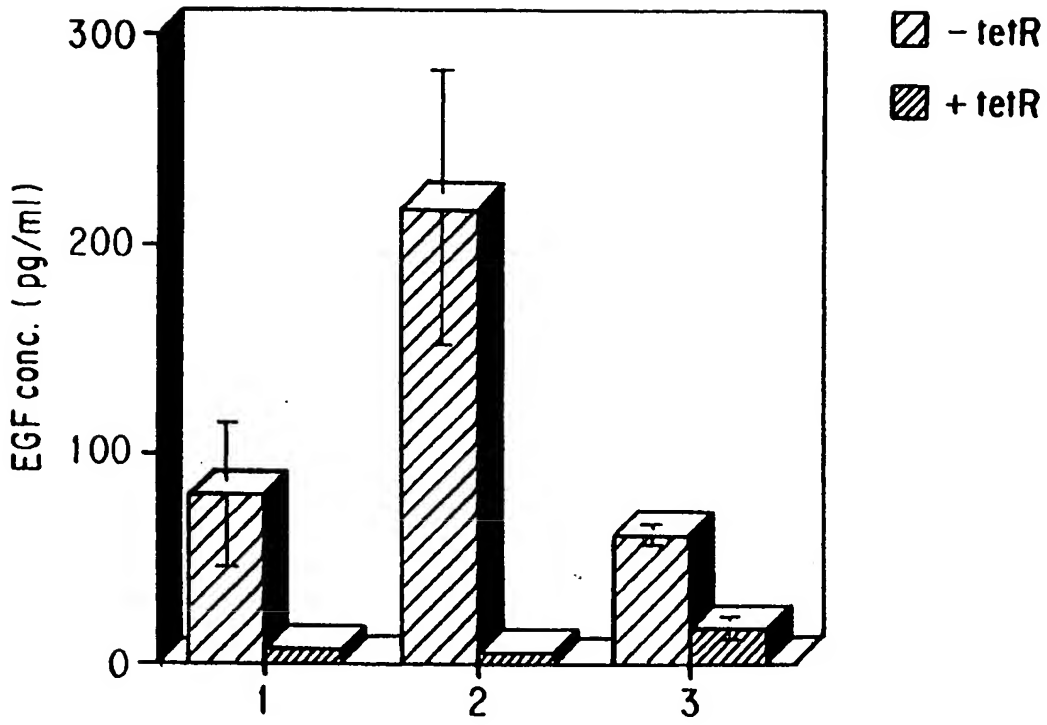

300
200
100
0
EGF conc. (pg/ml)
1
2
3
Days after in vivo gene transfer
- tetR
+ tetR

FIG. 5

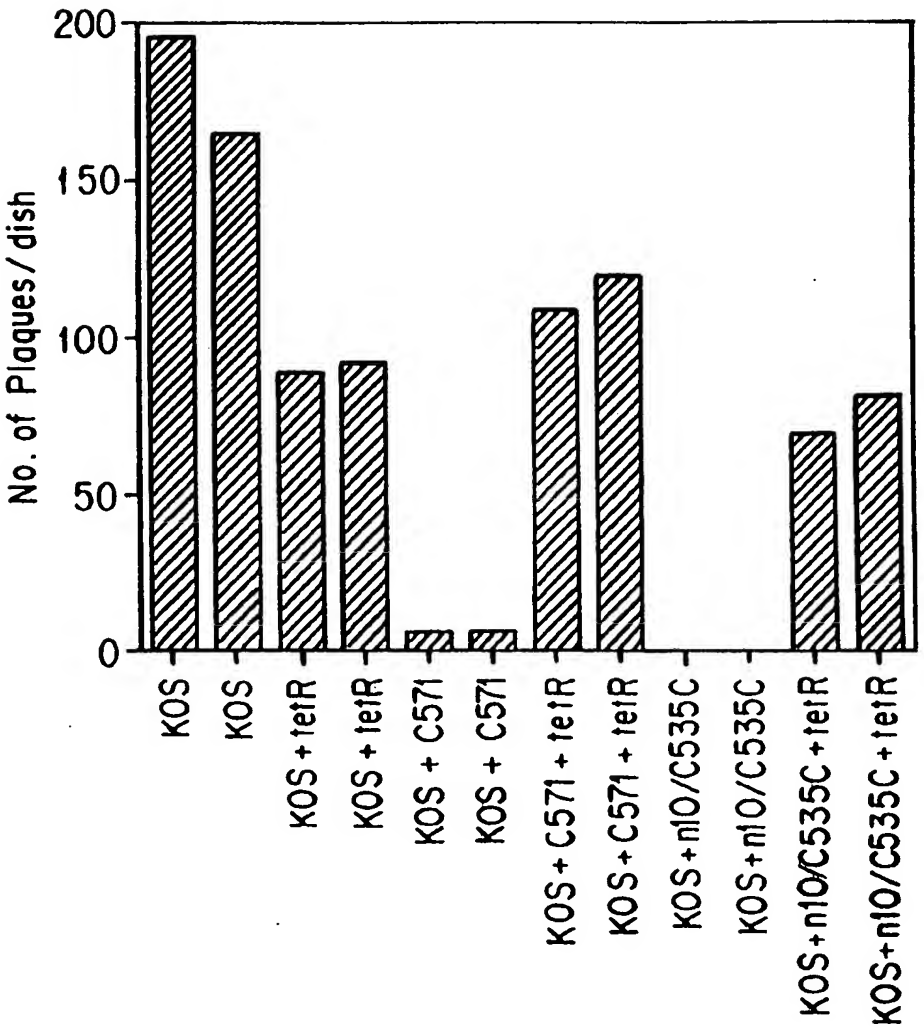

No. of Plaques/dish
200
150
100
50
0
KOS
KOS
KOS + tetR
KOS + tetR
KOS + C571
KOS + C571
KOS + C571 + tetR
KOS + C571 + tetR
KOS + n10/C535C
KOS + n10/C535C
KOS + n10/C535C + tetR
KOS + n10/C535C + tetR

FIG. 6

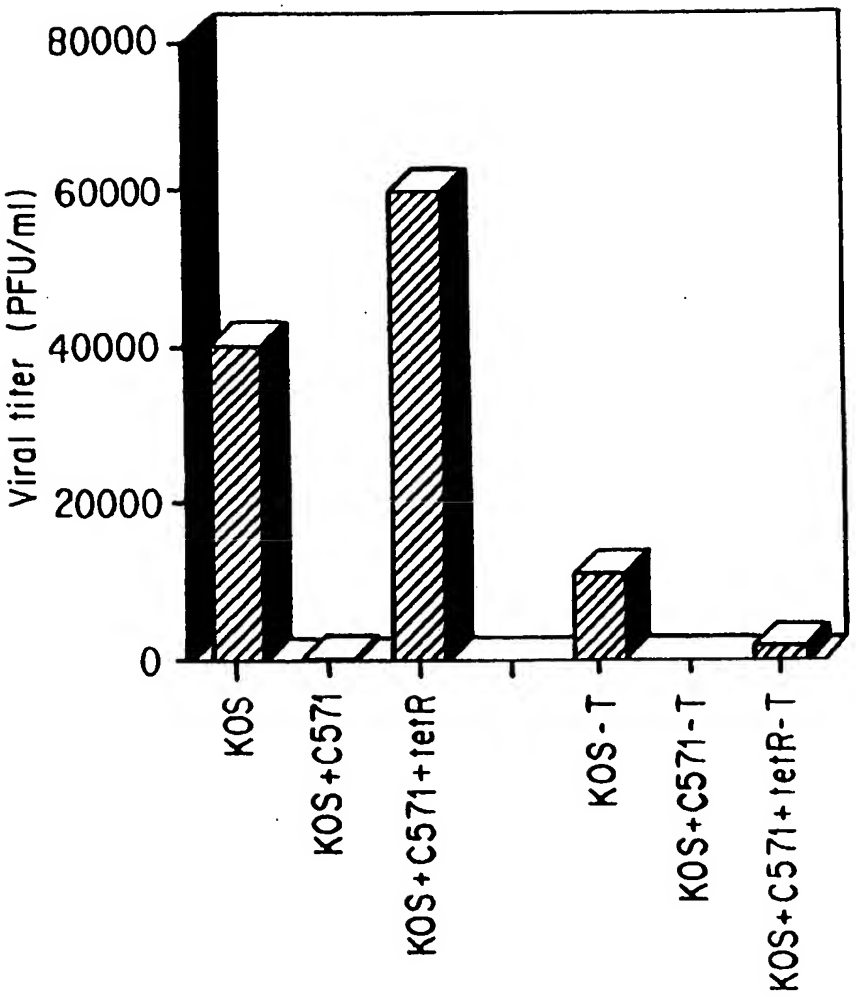

80000
60000
40000
20000
0
Viral titer (PFU/ml)
KOS
KOS+C571
KOS+C571+tetR
KOS-T
KOS+C571-T
KOS+C571+tetR-T

FIG. 7

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/10907

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/85 C12N15/86 A01K67/027 A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N A01K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 94 04672 A (DNX CORP ;BYRNE GUERARD (US)) 3 March 1994	1,2,4-8, 10-12, 14,15
Y	see page 25, line 23 - line 26; claims 2,5,7,10,14; figure 2	3,9,13, 16-25, 30-36
Y	WO 97 20463 A (GEN HOSPITAL CORP) 12 June 1997 see claims 10,24,31	3,9,13, 16-25, 30-36

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 August 1998	09/09/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Smalt, R

3

Form PCT/ISA/210 (second sheet) (July

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/10907

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.-
A	GATZ, C. ET AL.: "Regulation of a modified CaMV 35S promoter by the Tn10-encoded Tet repressor in transgenic tobacco" MOLECULAR AND GENERAL GENETICS, vol. 227, 1991, pages 229-37, XP002075578 see abstract; figure 1 see page 235, right-hand column, line 10 - line 15 ---	
A	WIRTZ, E. ET AL.: "Inducible gene expression in Trypanosomes mediated by a prokaryotic repressor." SCIENCE, vol. 268, 26 May 1995, pages 1179-83, XP002075579 see the whole document ---	2,8,18, 27
A	WO 93 01301 A (PENN STATE RES FOUND) 21 January 1993 see the whole document ---	26-29
A	WO 93 00446 A (GENELABS INC) 7 January 1993 see page 8, line 33 - line 35 see page 50, line 19 - line 23 ---	26-29
A	GATZ, C. ET AL.: "Stringent repression and homogenous de-repression by tetracycline of a modified CaMV 35S promoter in intact transgenic tobacco plants" THE PLANT JOURNAL, vol. 2, no. 3, 1992, pages 397-404, XP002075580 see abstract see page 398, left-hand column; figure 1A ---	
A	GOSSEN, M. ET AL.: "Control of gene activity in higher eukaryotic cells by prokaryotic regulatory elements" TRENDS IN BIOCHEMICAL SCIENCES, vol. 18, December 1993, pages 471-475, XP002075581 see page 473, middle column, paragraph 2 - right-hand column, paragraph 1 -----	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 98/ 10907

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Remark: Although claims 26-36 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Ir. ational Application No
PCT/US 98/10907

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9404672 A	03-03-1994	AU 683227 B	06-11-1997
		AU 5099393 A	15-03-1994
		CA 2143326 A	03-03-1994
		EP 0665883 A	09-08-1995
WO 9720463 A	12-06-1997	NONE	
WO 9301301 A	21-01-1993	US 5686601 A	11-11-1997
WO 9300446 A	07-01-1993	AT 165397 T	15-05-1998
		AU 655839 B	12-01-1995
		AU 2297892 A	25-01-1993
		CA 2112130 A,C	07-01-1993
		DE 69225227 D	28-05-1998
		EP 0593618 A	27-04-1994
		EP 0823486 A	11-02-1998
		JP 7500491 T	19-01-1995
		US 5306619 A	26-04-1994
		US 5726014 A	10-03-1998
		US 5578444 A	26-11-1996
		US 5738990 A	14-04-1998
		US 5744131 A	28-04-1998
		US 5716780 A	10-02-1998
		US 5693463 A	02-12-1997

EXHIBIT 3

Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate

Sayda M.Elbashir, Javier Martinez, Agnieszka Patkaniowska, Winfried Lendeckel and Thomas Tuschl[1]

Department of Cellular Biochemistry, Max-Planck-Institute for Biophysical Chemistry, Am Fassberg 11, D-37077 Göttingen, Germany

[1]Corresponding author
e-mail: ttuschl@mpibpc.gwdg.de

S.M.Elbashir and J.Martinez contributed equally to this work

Duplexes of 21–23 nucleotide (nt) RNAs are the sequence-specific mediators of RNA interference (RNAi) and post-transcriptional gene silencing (PTGS). Synthetic, short interfering RNAs (siRNAs) were examined in *Drosophila melanogaster* embryo lysate for their requirements regarding length, structure, chemical composition and sequence in order to mediate efficient RNAi. Duplexes of 21 nt siRNAs with 2 nt 3′ overhangs were the most efficient triggers of sequence-specific mRNA degradation. Substitution of one or both siRNA strands by 2′-deoxy or 2′-*O*-methyl oligonucleotides abolished RNAi, although multiple 2′-deoxynucleotide substitutions at the 3′ end of siRNAs were tolerated. The target recognition process is highly sequence specific, but not all positions of a siRNA contribute equally to target recognition; mismatches in the centre of the siRNA duplex prevent target RNA cleavage. The position of the cleavage site in the target RNA is defined by the 5′ end of the guide siRNA rather than its 3′ end. These results provide a rational basis for the design of siRNAs in future gene targeting experiments.

Keywords: PTGS/RNA interference/small interfering RNA

Introduction

Post-transcriptional gene silencing (PTGS) mediated by double-stranded (ds) RNA represents an evolutionarily conserved cellular defence mechanism for controlling the expression of alien genes in protists, filamentous fungi, plants and animals (Fire, 1999; Bass, 2000; Cogoni and Macino, 2000; Carthew, 2001; Hammond *et al.*, 2001b; Sharp, 2001; Tuschl, 2001; Voinnet, 2001; Waterhouse *et al.*, 2001). It is believed that random integration of alien genes (such as transposons) or viral infection causes production of dsRNA, which activates sequence-specific degradation of homologous single-stranded mRNA or viral genomic RNA, thereby preventing expression or replication of the foreign genetic material. The dsRNA is used as the guide RNA in this sequence-specific RNA degradation process. In some cases, dsRNA may also be involved in amplification of the silencing signal important for systemic spread (Palauqui *et al.*, 1997; Voinnet *et al.*, 1998) or long-term maintenance of silencing (Dalmay *et al.*, 2000; Mourrain *et al.*, 2000; Smardon *et al.*, 2000). In animals, the dsRNA-triggered silencing effect is referred to as RNA interference (RNAi; Fire *et al.*, 1998).

One important feature of the mechanism of RNAi is the processing of long dsRNAs into duplexes of 21–25 nucleotide (nt) RNAs. These short RNA products were first detected in plant tissues that exhibited transgene- or virus-induced PTGS (Hamilton and Baulcombe, 1999), but were also found later in fly embryos and worms injected with long dsRNAs (Parrish *et al.*, 2000; Yang *et al.*, 2000) or in extracts from *Drosophila melanogaster* Schneider-2 (S2) cells that were transfected with dsRNA (Hammond *et al.*, 2000). The processing reaction of long dsRNAs to 21–23 nt RNAs was first recapitulated *in vitro*, in extracts prepared from *D.melanogaster* embryos (Zamore *et al.*, 2000) and later in extracts from S2 cells (Bernstein *et al.*, 2001). In the embryo lysate, it was observed that the target mRNA was cleaved in ~21 nt intervals (Zamore *et al.*, 2000) and that synthetic 21 and 22 nt RNA duplexes added to the lysate were able to guide efficient sequence-specific mRNA degradation, while duplexes of 30 bp dsRNA were inactive (Elbashir *et al.*, 2001b). The 21 nt RNA products were therefore named small interfering RNAs or silencing RNAs (siRNAs).

A ribonuclease III enzyme, dicer, is required for processing of long dsRNA into siRNA duplexes (Bernstein *et al.*, 2001). It was recently shown that dicer has an additional cellular function and is also required for excision of 21 and 22 nt small temporal RNAs (stRNAs) from ~70 nt stable stem–loop precursors (Grishok *et al.*, 2001; Hutvágner *et al.*, 2001). These tiny expressed RNA molecules are important regulators of developmental timing and control the translation of downstream regulatory genes (Ambros, 2000; Moss, 2000; Pasquinelli *et al.*, 2000). stRNAs are different from siRNAs in that the target mRNA is not degraded during silencing (Wightman *et al.*, 1993; Olsen and Ambros, 1999) and they are single stranded (Reinhart *et al.*, 2000), while siRNAs are believed to be double stranded (Elbashir *et al.*, 2001b; Hutvágner *et al.*, 2001).

In RNAi, a siRNA-containing endonuclease complex cleaves a single-stranded target RNA in the middle of the region complementary to the 21 nt guide siRNA of the siRNA duplex (Elbashir *et al.*, 2001b). This cleavage site is one helical turn displaced from the cleavage site that produced the siRNA from long dsRNA, suggesting dramatic conformational and/or compositional changes after processing of long dsRNA to 21 nt siRNA duplexes. The target RNA cleavage products are rapidly degraded because they either lack the stabilizing cap or poly(A) tail. A protein component of the ~500 kDa endonuclease or

RNA-induced silencing complex (RISC) was recently identified and is a member of the argonaute family of proteins (Hammond *et al.*, 2001a); however, it is currently unclear whether dicer is required for RISC activity.

It is also unknown whether RISC contains single- or double-stranded siRNAs. By analogy to stRNA excision, it may be envisaged that only one of the strands of a siRNA duplex is incorporated into a catalytic siRNP, but because of the symmetry of the siRNA duplex, two approximately equal populations of sense and antisense strand-containing catalytic siRNPs are produced. Synthetic siRNA duplexes cleaved sense as well as antisense target RNAs in the middle of the region covered by the siRNA duplex in *D.melanogaster* lysate (Elbashir *et al.*, 2001b). However, longer dsRNAs did not produce symmetric sense and antisense target RNA cleavage sites in embryo lysate (Elbashir *et al.*, 2001b), suggesting that the direction of processing of long dsRNA defined which of the strands of the resulting siRNA duplex could be used for guiding target degradation. Some protein, involved in the production of the 21 nt siRNA duplexes, may be deposited on the siRNA duplex to mark the strand that is going to be used for guiding target RNA cleavage.

Despite the lack of profound mechanistic understanding, RNAi has rapidly developed into an important tool for reverse genetics and has been widely applied in *Caenorhabditis elegans* (Fraser *et al.*, 2000; Gönczy *et al.*, 2000; Piano *et al.*, 2000; Maeda *et al.*, 2001), as well as in insects (see references in Lam and Thummel, 2000) and insect cell lines (Clemens *et al.*, 2000; Hammond *et al.*, 2000; Ui-Tei *et al.*, 2000). RNAi has also been shown to occur in a variety of vertebrates by targeting of mRNAs important for embryonic development. In differentiated mammalian cells, dsRNAs with >30 bp generally activate the interferon response, which leads to a global shut-off in protein synthesis as well as non-specific mRNA degradation (Stark *et al.*, 1998). This unspecific response to long dsRNAs can be bypassed using 21 nt siRNA duplexes, resulting in specific knock-down of the expression of the targeted gene (Elbashir *et al.*, 2001a; Hutvágner *et al.*, 2001), providing a new method for analysis of mammalian gene function in cultured cells.

Here we describe the results of a systematic analysis of the length, secondary structure, sugar backbone and sequence specificity of siRNA duplexes for RNAi, using the established *D.melanogaster* embryo *in vitro* system. The most potent siRNA duplexes are 21 nt long, comprising a 19 nt base-paired sequence with 2 nt 3′-overhanging ends. The 5′ end of the target-complementary siRNA strand (guide siRNA) sets the ruler to define the position of target RNA cleavage. Furthermore, we find that target recognition is extremely specific, as even single nucleotide mismatches between the siRNA duplex and the target mRNA abolish interference. These results provide a rational basis for the design of siRNAs for future gene targeting experiments.

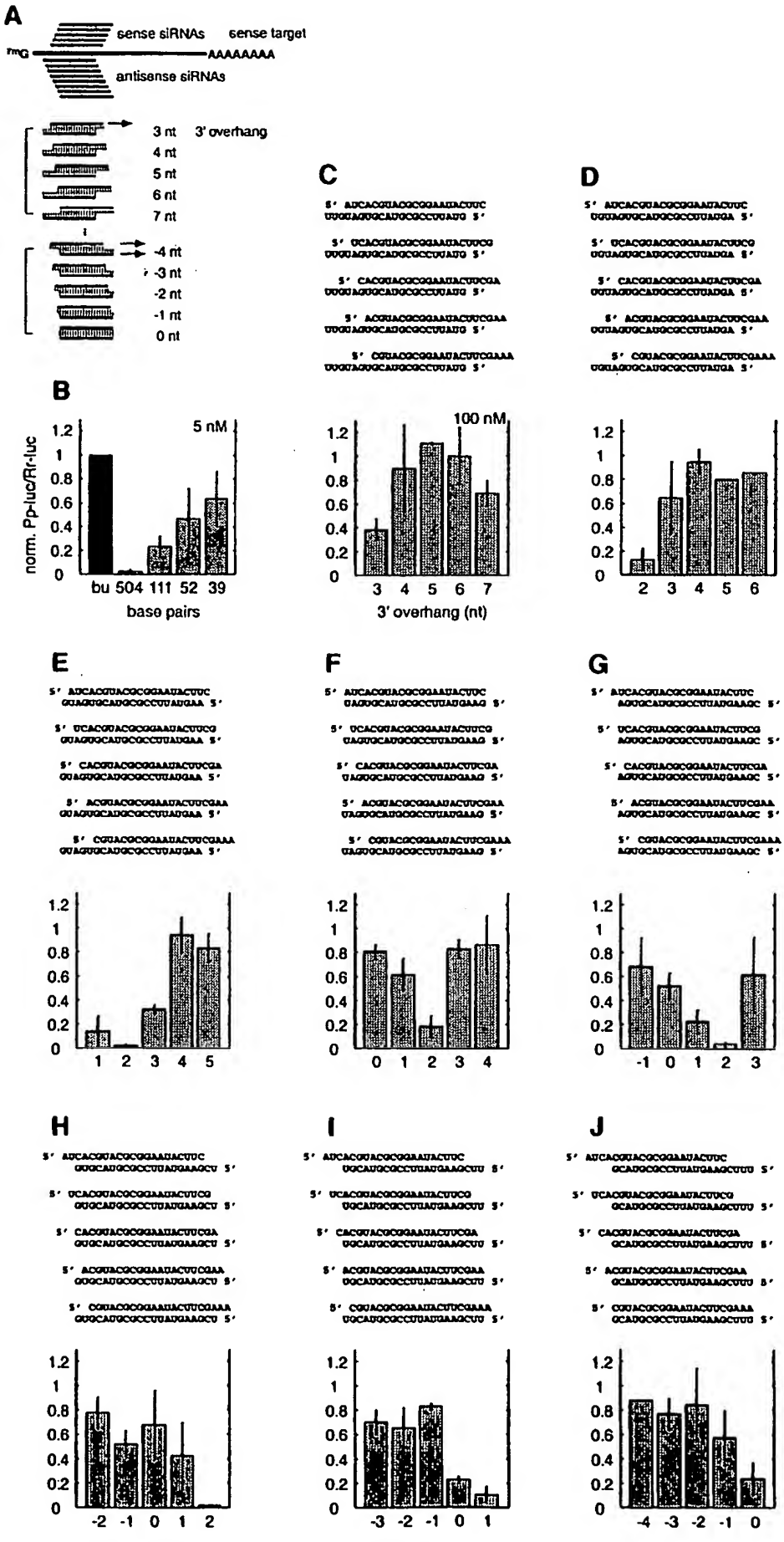
Results

Variation of the 3′ overhang in duplexes of 21 nt siRNAs

We reported previously that two or three unpaired nucleotides at the 3′ end of siRNA duplexes were more efficient in target RNA degradation than blunt-ended duplexes (Elbashir *et al.*, 2001b). To perform a more comprehensive analysis of the function of the terminal nucleotides, we synthesized five 21 nt sense siRNAs, each displaced by one nucleotide relative to the target RNA, and eight 21 nt antisense siRNAs, each displaced by one nucleotide relative to the target (Figure 1A). By combining these sense and antisense siRNAs, a series of eight siRNA duplexes with symmetric overhanging ends were generated spanning a range from 7 nt 3′ overhang to 4 nt 5′ overhang. The interference was measured using the dual luciferase assay system (Tuschl *et al.*, 1999; Zamore *et al.*, 2000). siRNA duplexes were directed against firefly luciferase mRNA and sea pansy luciferase mRNA was used as internal control. The luminescence ratio of target to control luciferase activity was determined in the presence of siRNA duplex and was normalized to that observed in its absence. For comparison, the interference ratios of long dsRNAs (39–504 bp) are shown in Figure 1B (Elbashir *et al.*, 2001b). The interference ratios were determined at concentrations of 5 nM for long dsRNAs (Figure 1A) and at 100 nM for siRNA duplexes (Figure 1C–J). The 100 nM concentration of siRNAs was chosen because complete processing of 5 nM 504 bp dsRNA would result in 120 nM total siRNA duplexes.

The ability of 21 nt siRNA duplexes to mediate RNAi is dependent on the number of overhanging nucleotides or base pairs formed. Duplexes with 4–6 3′-overhanging nucleotides were unable to mediate RNAi (Figure 1C–F), as were duplexes with two or more 5′-overhanging nucleotides (Figure 1G–J). The duplexes with 2 nt 3′ overhangs were most efficient in mediating RNA interference, although the efficiency of silencing was also sequence dependent and up to 12-fold differences were observed for different siRNA duplexes with 2 nt 3′ overhangs (compare Figure 1D–H). Duplexes with blunted ends, 1 nt 5′ overhang or 1–3 nt 3′ overhangs were sometimes functional and sometimes completely inactive. The small silencing effect observed for the siRNA duplex with 7 nt 3′ overhang (Figure 1C) may be due to an antisense effect of the long 3′ overhang rather than to

Fig. 1. Variation of the 3′ overhang of duplexes of 21 nt siRNAs. (**A**) Outline of the experimental strategy. The capped and polyadenylated sense target mRNA is depicted and the relative positions of sense and antisense siRNAs are shown. Eight series of duplexes according to the eight different antisense strands were prepared. The siRNA sequences and the number of overhanging nucleotides were changed in 1 nt steps. (**B**) Normalized relative luminescence of target luciferase (*Photinus pyralis*, Pp-luc) to control luciferase (*Renilla reniformis*, Rr-luc) in *D.melanogaster* embryo lysate in the presence of 5 nM blunt-ended dsRNAs (Elbashir *et al.*, 2001b). The luminescence ratios determined in the presence of dsRNA were normalized to the ratio obtained for a buffer control (bu; black bar). Normalized ratios less than 1 indicate specific interference. (**C–J**) Normalized interference ratios for eight series of 21 nt siRNA duplexes. The sequences of siRNA duplexes are depicted above the bar graphs. Each part shows the interference ratio for a set of duplexes formed with a given antisense guide siRNA and five different sense siRNAs. The number of overhanging nucleotides (3′ overhang, positive numbers; 5′ overhang, negative numbers) is indicated on the *x*-axis. Data points were averaged from at least three independent experiments. Error bars represent standard deviations.


A
sense siRNAs
sense target
m7G
AAAAAAAA
antisense siRNAs
3 nt
3' overhang
4 nt
5 nt
6 nt
7 nt
-4 nt
-3 nt
-2 nt
-1 nt
0 nt
B
5 nM
norm. Pp-luc/Rr-luc
bu 504 111 52 39
base pairs
C
100 nM
3 4 5 6 7
3' overhang (nt)
D
2 3 4 5 6
E
1 2 3 4 5
F
0 1 2 3 4
G
-1 0 1 2 3
H
-2 -1 0 1 2
I
-3 -2 -1 0 1
J
-4 -3 -2 -1 0

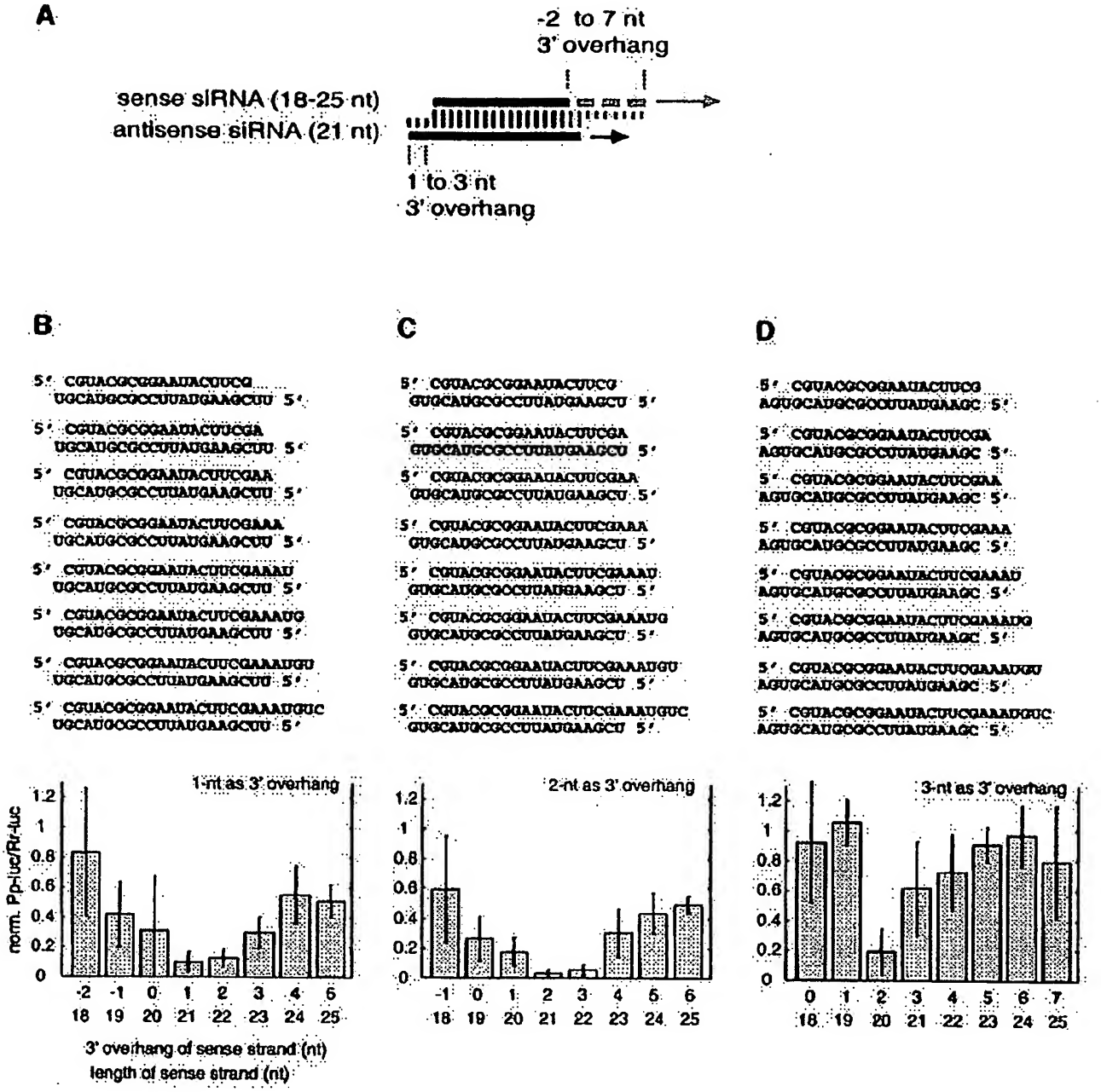


Fig. 2. Variation of the length of the sense strand of siRNA duplexes. (A) Representation of the experiment. Three 21 nt antisense strands were paired with eight sense siRNAs. The siRNAs were changed in length at their 3′ end. The 3′ overhang of the antisense siRNA was 1 nt (B), 2 nt (C) or 3 nt (D), while the sense siRNA overhang was varied for each series. The sequences of the siRNA duplexes and the corresponding interference ratios are indicated.

RNAi. Comparison of the efficiency of RNAi between long dsRNAs (Figure 1B) and the most effective 21 nt siRNA duplexes (Figure 1E, G and H) indicates that a single siRNA duplex at 100 nM concentration can be as effective as 5 nM 504 bp dsRNA.

Length variation of the sense siRNA paired to an invariant 21 nt antisense siRNA

In order to investigate the effect of the length of siRNAs on RNAi, we prepared three series of siRNA duplexes, combining three 21 nt antisense strands with eight 18–25 nt sense strands. The 3′ overhang of the antisense siRNA was fixed to 1, 2 or 3 nt in each siRNA duplex series, while the sense siRNA was varied at its 3′ end (Figure 2A). Independently of the length of the sense siRNA, we found that duplexes with 2 nt 3′ overhang of antisense siRNA (Figure 2C) were more active than those with 1 or 3 nt 3′ overhang (Figure 2B and D). In the first series, with 1 nt 3′ overhang of antisense siRNA, duplexes with 21 and 22 nt sense siRNAs, carrying a 1 and 2 nt 3′ overhang of sense siRNA, respectively, were most active. Duplexes with 19–25 nt sense siRNAs were also able to mediate RNA, but to a lesser extent. Similarly, in the second series, with 2 nt overhang of antisense siRNA, the 21 nt siRNA duplex with 2 nt 3′ overhang was most active and any other combination with the 18–25 nt sense siRNAs was active to

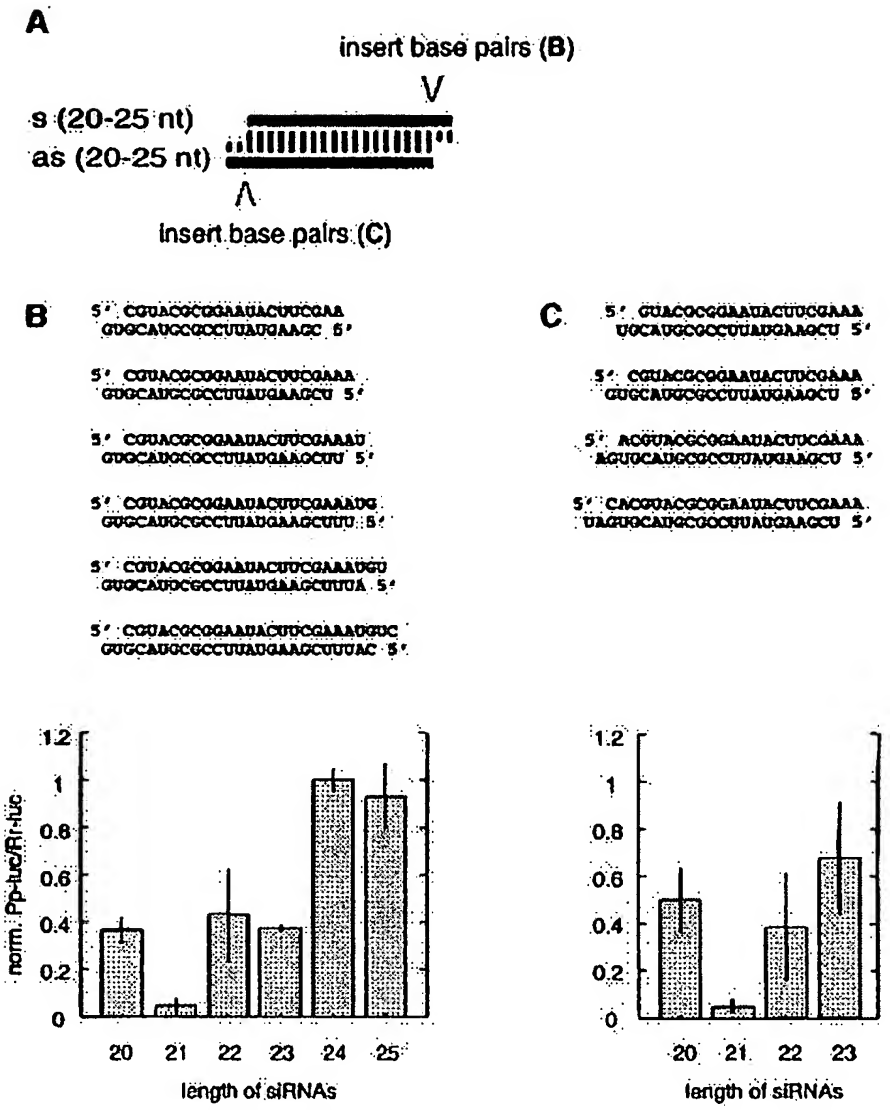


Fig. 3. Variation of the length of siRNA duplexes with preserved 2 nt 3′ overhangs. (A) Graphic representation of the experiment. The 21 nt siRNA duplex is identical in sequence to the one shown in Figures 1H and 2C. The siRNA duplexes were extended to the 3′ side of the sense siRNA (B) or the 5′ side of the sense siRNA (C). The siRNA duplex sequences and the respective interference ratios are indicated.

a significant degree. In the last series, with 3 nt antisense siRNA 3′ overhang, only the duplex with a 20 nt sense siRNA and 2 nt sense 3′ overhang was able to reduce target RNA expression. Together, these results indicate that the length of the siRNA as well as the length of the 3′ overhang are important, and that duplexes of 21 nt siRNAs with 2 nt 3′ overhang are optimal for RNAi.

Length variation of siRNA duplexes with a constant 2 nt 3′ overhang

We then examined the effect of simultaneously changing the length of both siRNA strands by maintaining symmetrical 2 nt 3′ overhangs (Figure 3A). Two series of siRNA duplexes were prepared, including the 21 nt siRNA duplex of Figure 1H as reference. The length of the duplexes was varied between 20 and 25 bp by extending the base-paired segment at the 3′ end of the sense siRNA (Figure 3B) or at the 3′ end of the antisense siRNA (Figure 3C).

Duplexes of 20–23 bp caused specific repression of target luciferase activity, but the 21 nt siRNA duplex was at least 8-fold more efficient than any of the other duplexes. siRNA duplexes of 24 and 25 nt did not result in any detectable interference. Sequence-specific effects were minor as variations on both ends of the duplex produced similar effects.

2′-deoxy- and 2′-O-methyl-modified siRNA duplexes

To assess the importance of the siRNA ribose residues for RNAi, duplexes with 21 nt siRNAs and 2 nt 3′ overhangs with 2′-deoxy- or 2′-*O*-methyl-modified strands were examined (Figure 4). Substitution of the 2 nt 3′ overhangs by 2′-deoxynucleotides had no effect and even the replacement of two additional ribonucleotides by 2′-deoxyribonucleotides adjacent to the overhangs in the paired region produced significantly active siRNAs. Thus, 8 out of 42 nt of a siRNA duplex were replaced by DNA

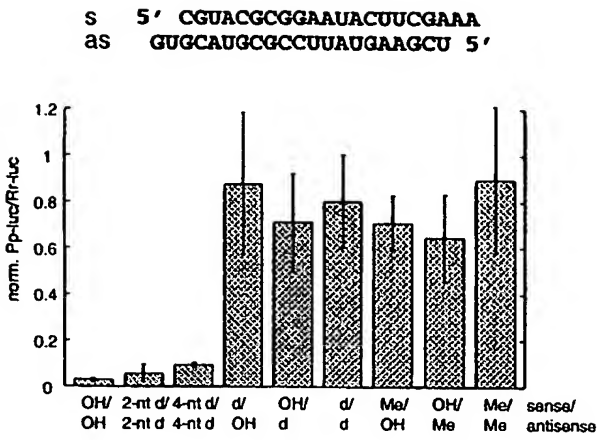


Fig. 4. Substitution of the 2′-hydroxyl groups of the siRNA ribose residues. The 2′-hydroxyl groups (OH) in the strands of siRNA duplexes were replaced by 2′-deoxy (d) or 2′-*O*-methyl (Me). 2 and 4 nt 2′-deoxy substitutions at the 3′ ends are indicated as 2- and 4-nt d, respectively. Uridine residues were replaced by 2′-deoxythymidine.

residues without loss of activity. Complete substitution of one or both siRNA strands by 2′-deoxy residues, however, abolished RNAi, as did complete substitution by 2′-*O*-methyl residues.

Definition of target RNA cleavage sites

Target RNA cleavage positions were previously determined for 22 nt siRNA duplexes and for a 21 and 22 nt duplex (Elbashir *et al.*, 2001b). The position of target RNA cleavage was located in the centre of the region covered by the siRNA duplex, 11 or 12 nt downstream of the first nucleotide that was complementary to the 21 or 22 nt siRNA guide sequence. Five distinct 21 nt siRNA duplexes with 2 nt 3′ overhang (Figure 5A) were incubated with 5′ cap-labelled sense or antisense target RNA in *D.melanogaster* lysate (Tuschl *et al.*, 1999; Zamore *et al.*, 2000). The 5′ cleavage products were resolved on sequencing gels (Figure 5B). The amount of sense target RNA cleaved correlated with the efficiency of siRNA duplexes determined in the translation-based assay, and siRNA duplexes 1, 2 and 4 (Figures 5B, 1E, G and H) cleaved target RNA faster than duplexes 3 and 5 (Figures 5B, 1D and F). Notably, the sum of radioactivity of the 5′ cleavage product and the input target RNA were not constant over time and the 5′ cleavage products did not accumulate. Presumably, the cleavage products, once released from the siRNA–endonuclease complex, were rapidly degraded due to the lack of either the poly(A) tail or the 5′ cap.

The cleavage sites for both sense and antisense target RNAs were located in the middle of the region spanned by the siRNA duplexes. The cleavage sites for each target produced by the five different duplexes varied by 1 nt according to the 1 nt displacement of the duplexes along the target sequences. The targets were cleaved precisely 11 nt downstream of the target position complementary to the 3′-most nucleotide of the sequence-complementary guide siRNA (Figure 5).

In order to determine whether the 5′ or the 3′ end of the guide siRNA sets the ruler for target RNA cleavage, we devised the experimental strategy outlined in Figure 6A

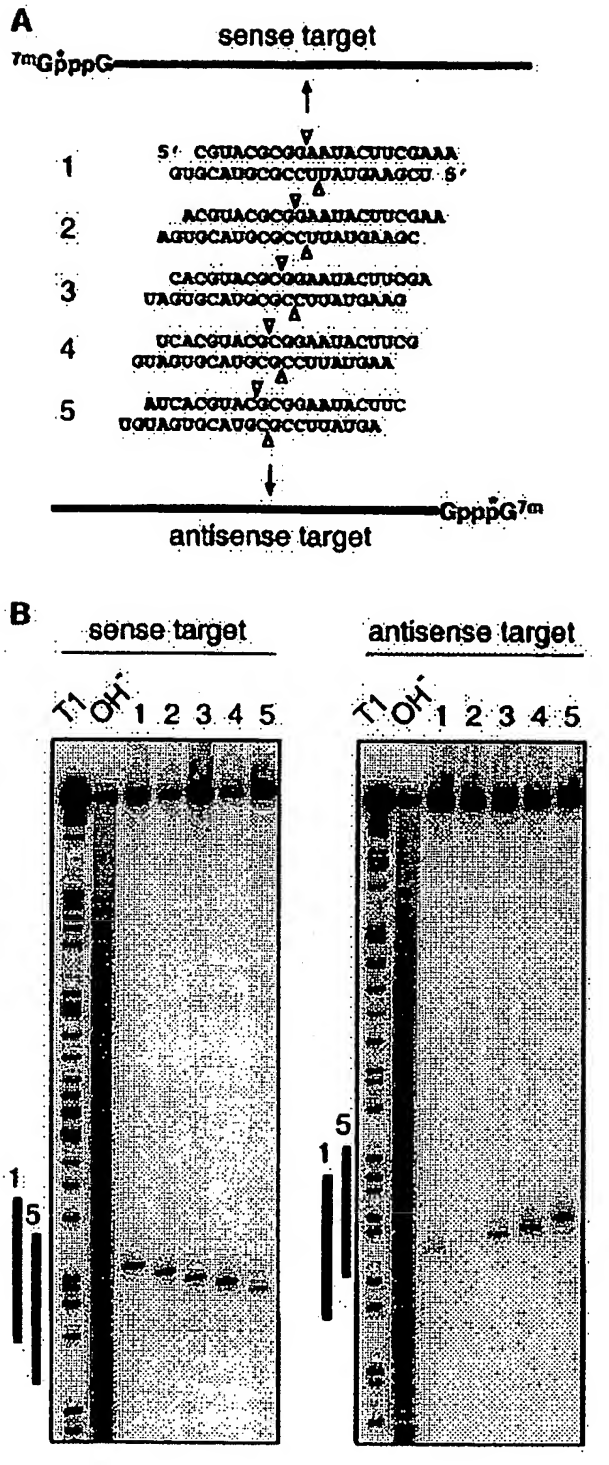


Fig. 5. Mapping of sense and antisense target RNA cleavage by 21 nt siRNA duplexes with 2 nt 3′ overhangs. (A) Representation of ^{32}P (asterisk) cap-labelled sense and antisense target RNAs and siRNA duplexes. The position of sense and antisense target RNA cleavage is indicated by triangles on top and below the siRNA duplexes, respectively. (B) Mapping of target RNA cleavage sites. After 2 h incubation of 10 nM target RNA with 100 nM siRNA duplex in *D.melanogaster* embryo lysate, the 5′ cap-labelled substrate and the 5′ cleavage products were resolved on 6% sequencing gels. Length markers were generated by partial RNase T1 digestion (T1) and partial alkaline hydrolysis (OH) of the target RNAs. The bold lines to the left of the images indicate the region covered by the siRNA strands 1 and 5 of the same orientation as the target.

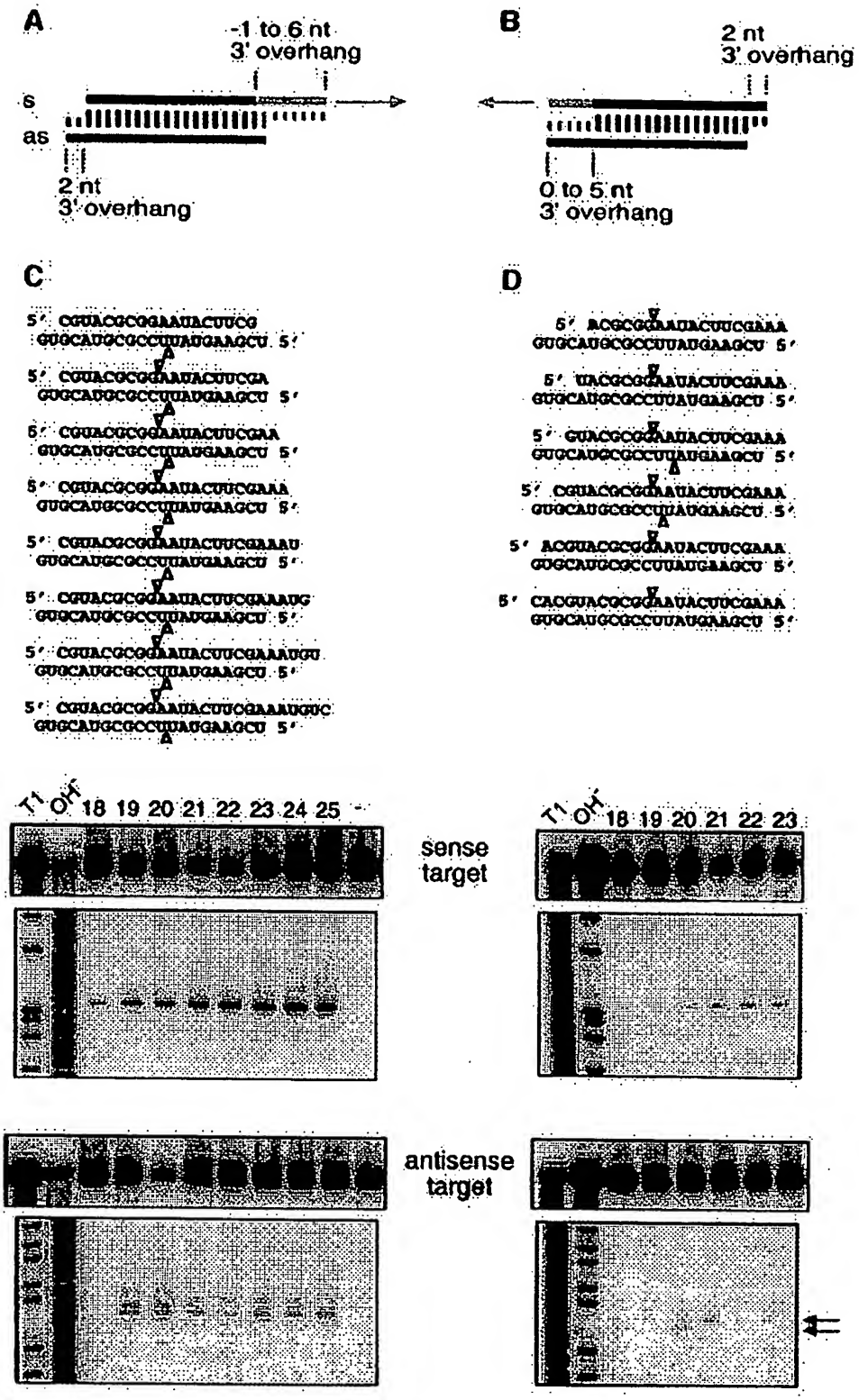


Fig. 6. The 5′ end of a guide siRNA defines the position of target RNA cleavage. (**A** and **B**) Representation of the experimental strategy. The antisense siRNA was the same in all siRNA duplexes, but the sense strand was varied between 18 and 25 nt by changing the 3′ end (A) or 18 and 23 nt by changing the 5′ end (B). The position of sense and antisense target RNA cleavage is indicated by triangles on top and below the siRNA duplexes, respectively. (**C** and **D**) Analysis of target RNA cleavage using cap-labelled sense (top) or antisense (bottom) target RNAs. The residual amount of targeted substrate and the cap-labelled 5′ cleavage products are shown. The sequences of the siRNA duplexes are indicated and the length of the sense siRNA strands is marked on top. The control lane, marked with a dash in (C), shows target RNA incubated in absence of siRNAs. Markers were as described in Figure 5. The arrows in (D), bottom, indicate the target RNA cleavage sites that differ by 1 nt.

and B. A 21 nt antisense siRNA, which was kept invariant for this study, was paired with sense siRNAs that were modified in length at either of their 5′ or 3′ ends. The position of sense and antisense target RNA cleavage was

determined as described above. Changes in the 3′ end of the sense siRNA, monitored for 1 nt 5′ overhang to 6 nt 3′ overhang, did not affect either the position of sense nor antisense target RNA cleavage (Figure 6C). Changes in the 5′ end of the sense siRNA did not affect the sense target RNA cleavage (Figure 6D, top), as expected, because the antisense siRNA was unchanged. The residual amount of uncleaved sense target RNA (Figure 6C and D top) correlated with the efficiency of siRNA duplexes determined in translation-based assays (Figure 2C and data not shown), but did not correlate with the amount of detected cleavage product. Accumulation of cleavage products was more pronounced for the longer and less efficient siRNA duplexes, suggesting that product release may have become rate limiting. Because the antisense siRNA was kept unchanged while the sense siRNA was varied, an alteration in product release implies a role of the sense siRNA strand in the target RNA degradation process.

Changes in the 5′ end of the sense siRNA, in contrast to its 3′ end, strongly affected antisense target RNA cleavage (Figure 6D, bottom). The antisense target was only cleaved when the sense siRNA was 20 or 21 nt in size. The position of cleavage differed by 1 nt, suggesting that the 5′ end of the target-recognizing siRNA sets the ruler for target RNA cleavage. This position is located between nucleotide 10 and 11 when counting in an upstream direction from the target nucleotide paired to the 5′-most nucleotide of the guide siRNA (see also Figure 5A).

Sequence effects and 2′-deoxy substitutions in the 3′ overhang

The 2 nt 3′ overhang is critical for siRNA function. We wanted to know whether the sequence of the overhanging nucleotides contributes to target recognition or is only a feature required for reconstitution of the endonuclease complex (RISC or siRNP). We synthesized sense and antisense siRNAs with AA, CC, GG, UU and UG 3′ overhangs and included the 2′-deoxy modifications TdG and TT (T, 2′-deoxythymidine; dG, 2′-deoxyguanosine). The wild-type siRNAs contained AA in the sense 3′ overhang and UG in the antisense 3′ overhang (AA/UG). All siRNA duplexes were functional in the interference assay and reduced target expression at least 5-fold (Figure 7). The most efficient siRNA duplexes, which reduced target expression >10-fold, were of the sequence type NN/UG, NN/UU, NN/TdG and NN/TT (N, any nucleotide). siRNA duplexes with an antisense siRNA 3′ overhang of AA, CC or GG were less active by a factor of 2–4 when compared with the wild-type sequence UG or the mutant UU. This reduction in RNAi efficiency is likely to be due to the contribution of the penultimate 3′ nucleotide to sequence-specific target recognition, as the 3′-terminal nucleotide was changed from G to U without effect.

Changes in the sequence of the 3′ overhang of the sense siRNA did not reveal any sequence-dependent effects, which was not surprising because the sense siRNA is not expected to contribute to the sequence-specific recognition of the sense target mRNA.

Sequence specificity of target recognition

In order to examine the sequence specificity of target recognition, we introduced sequence changes into the

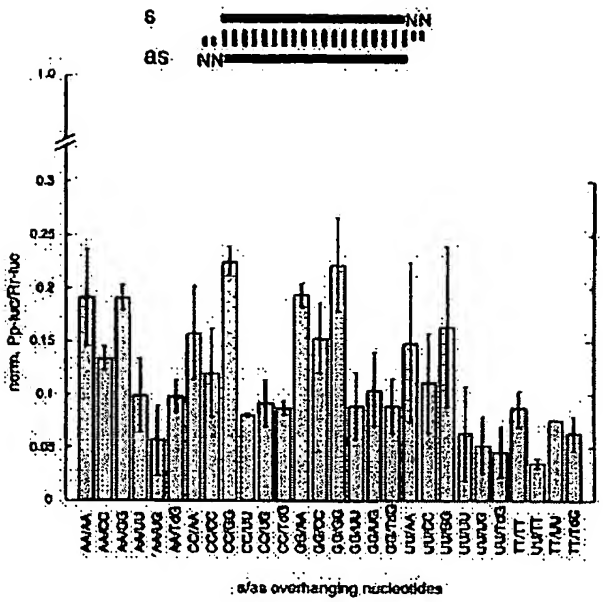


Fig. 7. Sequence variation of the 3′ overhang of siRNA duplexes. The 2 nt 3′ overhang (NN, in grey) was changed in sequence and composition as indicated (T, 2′-deoxythymidine; dG, 2′-deoxyguanosine; asterisk, wild-type siRNA duplex). Normalized interference ratios were determined as described in Figure 1. The wild-type sequence is the same as depicted in Figure 4.

paired segments of siRNA duplexes and determined the efficiency of silencing. Sequence changes were introduced by inverting short segments of 3 or 4 nt or inducing point mutations (Figure 8). The sequence changes in one siRNA strand were compensated for in the complementary siRNA strand to avoid perturbing the base-paired siRNA duplex structure. The sequence of all 2 nt 3′ overhangs was TT to reduce costs of synthesis. The TT/TT reference siRNA duplex was comparable in RNAi to the wild-type siRNA duplex AA/UG (Figure 7). The ability to mediate reporter mRNA destruction was quantified using the translation-based luminescence assay. Duplexes of siRNAs with inverted sequence segments showed dramatically reduced ability for targeting the firefly luciferase reporter (Figure 8). The sequence changes located between the 3′ end and the middle of the antisense siRNA completely abolished target RNA recognition, but mutations near the 5′ end of the antisense siRNA exhibited a small degree of silencing. Transversion of the AU base pair located directly opposite the predicted target RNA cleavage site or 1 nt further away from the predicted site prevented target RNA cleavage, therefore indicating that a single mutation within the centre of a siRNA duplex discriminates between mismatched targets.

Discussion

siRNAs are valuable reagents for inactivation of gene expression, not only in insect cells but also in mammalian cells, with a great potential for therapeutic application (Elbashir *et al.*, 2001a). We have systematically analysed the structural determinants of siRNA duplexes required to promote efficient target RNA degradation in *D.melanogaster* embryo lysate, thus providing rules for the design of most potent siRNA duplexes. A perfect siRNA duplex is able to silence gene expression with an

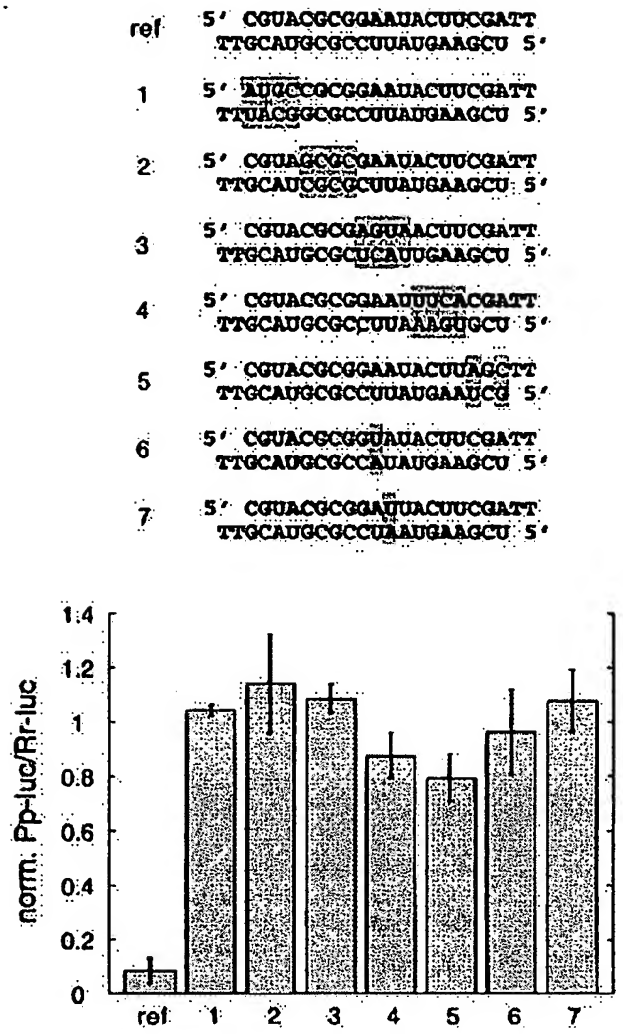


Fig. 8. Sequence specificity of target recognition. The sequences of the mismatched siRNA duplexes are shown, modified sequence segments or single nucleotides are shaded in grey. The reference duplex (ref) and the siRNA duplexes 1–7 contain 2′-deoxythymidine 2 nt overhangs. The silencing efficiency of the thymidine-modified reference duplex was comparable to the wild-type sequence (Figure 7). Normalized interference ratios were determined as described in Figure 1.

efficiency comparable to a 500 bp dsRNA, given that comparable quantities of total RNA are used.

The siRNA user guide

Efficiently silencing siRNA duplexes are composed of 21 nt sense and 21 nt antisense siRNAs and must be selected to form a 19 bp double helix with 2 nt 3′-overhanging ends. 2′-deoxy substitutions of the 2 nt 3′-overhanging ribonucleotides do not affect RNAi, but help to reduce the costs of RNA synthesis and may enhance RNase resistance of siRNA duplexes. More extensive 2′-deoxy or 2′-*O*-methyl modifications reduce the ability of siRNAs to mediate RNAi, probably by interfering with protein association for siRNP assembly.

Target recognition is a highly sequence-specific process, mediated by the siRNA complementary to the target. The 3′-most nucleotide of the guide siRNA does not contribute to the specificity of target recognition, while the penultimate nucleotide of the 3′ overhang affects target RNA cleavage and a mismatch reduces RNAi 2- to 4-fold. The 5′ end of the guide siRNA also appears more permissive for mismatched target RNA recognition when compared with the 3′ end. Nucleotides in the centre of the siRNA, located opposite to the target RNA cleavage site, are important specificity determinants and even single nucleotide changes reduce RNAi to undetectable levels. This suggests that siRNA duplexes may be able to discriminate mutant or polymorphic alleles in gene targeting experiments, which may become an important feature for future therapeutic developments.

Sense and antisense siRNAs, when associated with the protein components of the endonuclease complex or its commitment complex, were suggested to play distinct roles; the relative orientation of the siRNA duplex in this complex defines which strand can be used for target recognition (Elbashir *et al.*, 2001b). Synthetic siRNA duplexes with an equal number of overhanging nucleotides have dyad symmetry with respect to the double-helical structure, but not with respect to sequence. The association of siRNA duplexes with the RNAi proteins in the *D.melanogaster* lysate leads to the formation of two asymmetric complexes. In such hypothetical complexes, the chiral environment is distinct for sense and antisense siRNA, hence their function. The prediction obviously does not apply to palindromic siRNA sequences or to RNAi proteins that could associate as homodimers. To minimize sequence effects that may affect the ratio of sense- and antisense-targeting siRNPs, we suggest using siRNA sequences with identical 3′-overhanging sequences. We recommend adjusting the sequence of the overhang of the sense siRNA to that of the antisense 3′ overhang because the sense siRNA does not have a target in typical knock-down experiments. Asymmetry in the reconstitution of sense- and antisense-cleaving siRNPs could be, partially, responsible for the variation in RNAi efficiency observed for various 21 nt siRNA duplexes with 2 nt 3′ overhangs used in this study (Figure 1). Alternatively, the nucleotide sequence at the target site and/or the accessibility of the target RNA structure may be responsible for the variation in efficiency observed for these siRNA duplexes. It should be noted that all siRNAs used in this study are derived from a short region of one gene. Thus, it is more likely that differences in siRNA efficiency are a consequence of the primary sequences of the siRNAs and the respective target sites, rather than the secondary or tertiary structure of the targeted RNA.

Natural siRNAs versus synthetic siRNAs

In *D.melanogaster*, siRNA duplexes are produced *in vitro* and *in vivo* from long dsRNAs (Hammond *et al.*, 2000; Yang *et al.*, 2000; Zamore *et al.*, 2000). About 45% of these short RNAs are precisely 21 nt long, 28% are 22 nt long and a few percent are shorter or longer RNAs (Elbashir *et al.*, 2001b). This length distribution correlates with our finding that 21 nt siRNA duplexes are the most efficient mediators of mRNA degradation. Beside the length, the paired structure and overhang are also important. This structural feature may explain why siRNA duplexes isolated from the dsRNA processing reaction under denaturing conditions were less potent for RNAi than longer dsRNAs that were processed to siRNAs during the targeting reaction (Zamore *et al.*, 2000). Presumably,

denaturation followed by renaturation favoured the formation of the thermodynamically more stable, blunt-ended, but less active, siRNA duplexes. Isolation of siRNAs under native conditions does not reduce siRNA activity (Nykänen *et al.*, 2001).

Production of siRNAs from long dsRNA requires the RNase III enzyme dicer (Bernstein *et al.*, 2001). Dicer is a bidentate RNase III, which also contains an ATP-dependent RNA helicase domain and a PAZ domain, presumably important for dsRNA unwinding and mediation of protein–protein interactions, respectively (Cerutti *et al.*, 2000; Bernstein *et al.*, 2001). Dicer is evolutionarily conserved in worms, flies, plants, fungi and mammals (Matsuda *et al.*, 2000), and has a second cellular function important for the development of these organisms (Ray *et al.*, 1996; Jacobsen *et al.*, 1999; Grishok *et al.*, 2001; Hutvágner *et al.*, 2001; Knight and Bass, 2001). At present, it is uncertain whether dicer activity in species other than *D.melanogaster* produces siRNAs of predominantly 21 nt in length. The estimates of siRNA size vary in the literature between 21 and 25 nt (Hamilton and Baulcombe, 1999; Hammond *et al.*, 2000; Hutvágner *et al.*, 2000; Parrish *et al.*, 2000; Yang *et al.*, 2000; Zamore *et al.*, 2000; Elbashir *et al.*, 2001b).

In a recent study of the effect of siRNA length in mammalian cells (primary mouse embryonic fibroblasts, 293 and HeLa cells), duplexes of 21–27 nt siRNAs with 2 nt 3′ overhangs were directed against different co-transfected reporter genes (Caplen *et al.*, 2001). Duplexes of 22 and 23 nt siRNAs were found to be slightly more efficient in triggering sequence-specific gene silencing than 21 nt siRNA duplexes. In our hands, using the dual luciferase assay system in HeLa cells, 21 nt siRNA duplexes with 2 nt 3′ overhang are 2- to 3-fold more efficient than 20 or 22–25 nt siRNA duplexes (data not shown), therefore recapitulating the results obtained from the *D.melanogaster* biochemical system. In contrast to the *D.melanogaster* system, siRNA duplexes >23 nt in length are still triggering some RNAi in HeLa cells and also in *C.elegans* (Caplen *et al.*, 2001). However, it remains to be determined whether the RNA strands finally incorporated into the active endonuclease complex are of the initially provided length. It is possible that exonucleases present in *C.elegans* and mammalian cells trim longer siRNAs to their optimum length and that these exonucleases are absent from *D.melanogaster* lysate.

The functional anatomy of long dsRNAs as a trigger for RNAi was analysed previously in *C.elegans* (Parrish *et al.*, 2000). Activation of RNAi by injection of long dsRNA requires at least two steps: dsRNA processing by dicer RNase III and siRNP or RISC formation. Substitution of one of the strands of the long dsRNA by DNA abolished RNAi and even the substitution of C by dC or U by dT in only one of the strands caused a substantial decrease in RNAi. Because introduction of 2′-fluoro modifications into long RNA had no effect, it was suggested that an A-form double helical structure was important for triggering RNAi (Parrish *et al.*, 2000). We have been able to substitute eight ribose residues of a siRNA duplex by 2′-deoxyribose residues without substantial reduction of RNAi, although it should be noted the 2′-deoxy modifications were clustered at the 3′ end of the siRNAs, including the 2 nt 3′ overhangs. It is possible that the four 2′-deoxy modifications, which are located in the paired region at the end of the helix, do not affect the overall A-form helical structure and do not strongly compromise RISC formation. Complete modification of one or both siRNA strands by 2′-deoxyribose, however, abolished RNAi. Interestingly, substitution by 2′-*O*-methylribose, which adopts the ribose sugar pucker, also abolished RNAi, probably because methylation of the 2′-hydroxyls blocked hydrogen bond formation or introduced steric hindrance.

It was recently demonstrated that a 5′-phosphate on the target-complementary strand of a siRNA duplex is required for siRNA function and that ATP is used to maintain the 5′-phosphates of the siRNAs (Nykänen *et al.*, 2001). However, 5′-phosphorylation of fully 2′-deoxy- or 2′-*O*-methyl-modified siRNA strands was not able to restore RNAi (data not shown). Unmodified siRNA duplexes with free 5′-hydroxyls and 2 nt 3′ overhangs are readily phosphorylated in *D.melanogaster* embryo lysate (Nykänen *et al.*, 2001). In this respect, it should be noted that our reported RNAi efficiencies were determined by pre-incubating the siRNA duplexes for 15 min in *D.melanogaster* lysate before adding target and control mRNAs, thus providing sufficient time for 5′-phosphorylation of siRNA duplexes to occur. Comparison of the RNAi efficiencies of 5′-phosphorylated and 5′-non-phosphorylated siRNAs (for duplexes shown in Figures 1E, F and 2C) did not reveal any sizeable differences (data not shown).

Conclusions

We have performed an extensive analysis of the length, sequence and structure of siRNA duplexes in *D.melanogaster* embryo lysate. Duplexes of 21 nt siRNAs with 2 nt 3′ overhangs were shown to be the most efficient triggers of RNAi-based mRNA degradation. The target recognition is a highly sequence-specific process, although not all positions of a guide siRNA contribute equally to specificity. These results are important for the design of efficient siRNAs in order to silence genes in *D.melanogaster* and provide a basis for similar studies in other organisms.

Materials and methods

RNA preparation and RNAi assay

Chemical RNA synthesis, annealing and luciferase-based RNAi assays were performed as described previously (Tuschl *et al.*, 1999; Zamore *et al.*, 2000; Elbashir *et al.*, 2001b). Synthetic RNAs were gel purified after deprotection. The formation of siRNA duplexes was verified by agarose gel electrophoresis using 4% NuSieve GTG agarose (BMA, Rockland, ME) in 0.5× TBE buffer. All siRNA duplexes were directed against firefly luciferase and the luciferase mRNA sequence was derived from pGEM-luc (DDBJ/EMBL/GenBank accession No. X65316) as described (Tuschl *et al.*, 1999). The siRNA duplexes were incubated in a *D.melanogaster* RNAi/translation reaction for 15 min prior to addition of mRNAs. Translation-based RNAi assays were performed at least in triplicate.

For mapping of sense target RNA cleavage, a 177 nt transcript was generated, corresponding to the firefly luciferase sequence between positions 113 and 273 relative to the start codon, followed by the 17 nt complement of the SP6 promoter sequence (Elbashir *et al.*, 2001b). For mapping of antisense target RNA cleavage, a 166 nt transcript was produced from a template, which was amplified from plasmid sequence by PCR using the 5′ primer TAATACGACTCACTATAGAGCCCATATCGTTTCATA (T7 promoter underlined) and 3′ primer AGAG-

GATGGAACCGCTGG. The target sequence corresponds to the complement of the firefly luciferase sequence between positions 50 and 215 relative to the start codon. Guanylyl transferase labelling was performed as described previously (Zamore *et al.*, 2000). For mapping of target RNA cleavage, 100 nM siRNA duplex was incubated with 5–10 nM target RNA in *D.melanogaster* embryo lysate under standard conditions (Zamore *et al.*, 2000) for 2 h at 25°C. The reaction was stopped by the addition of 8 vols of proteinase K buffer [200 mM Tris–HCl pH 7.5, 25 mM EDTA, 300 mM NaCl, 2% (w/v) SDS]. Proteinase K (dissolved in water; Merck) was added to a final concentration of 0.6 mg/ml. The reactions were then incubated for 15 min at 65°C, extracted with phenol/chloroform/isoamyl alcohol (25:24:1) and precipitated with 3 vols of ethanol. Samples were loaded on 6% sequencing gels. Length standards were generated by partial RNase T1 digestion and partial base hydrolysis of the cap-labelled sense or antisense target RNAs.

Acknowledgements

We acknowledge Heike Taubner and Jutta Meyer for technical assistance, Phil Zamore for critical comments on the manuscript and H.Jäckle and R.Lührmann for support. This work was funded by a BMBF Biofuture grant, No. 0311856.

References

Ambros,V. (2000) Control of developmental timing in *Caenorhabditis elegans*. *Curr. Opin. Genet. Dev.*, **10**, 428–433.

Bass,B.L. (2000) Double-stranded RNA as a template for gene silencing. *Cell*, **101**, 235–238.

Bernstein,E., Caudy,A.A., Hammond,S.M. and Hannon,G.J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, **409**, 363–366.

Caplen,N.J., Parrish,S., Imani,F., Fire,A. and Morgan,R.A. (2001) Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems. *Proc. Natl Acad. Sci. USA*, **98**, 9742–9747.

Carthew,R.W. (2001) Gene silencing by double-stranded RNA. *Curr. Opin. Cell Biol.*, **13**, 244–248.

Cerutti,L., Mian,N. and Bateman,A. (2000) Domains in gene silencing and cell differentiation proteins: the novel PAZ domain and redefinition of the piwi domain. *Trends Biochem. Sci.*, **25**, 481–482.

Clemens,J.C., Worby,C.A., Simonson-Leff,N., Muda,M., Maehama,T., Hemmings,B.A. and Dixon,J.E. (2000) Use of double-stranded RNA interference in *Drosophila* cell lines to dissect signal transduction pathways. *Proc. Natl Acad. Sci. USA*, **97**, 6499–6503.

Cogoni,C. and Macino,G. (2000) Post-transcriptional gene silencing across kingdoms. *Curr. Opin. Genet. Dev.*, **10**, 638–643.

Dalmay,T., Hamilton,A., Rudd,S., Angell,S. and Baulcombe,D.C. (2000) An RNA-dependent RNA polymerase gene in *Arabidopsis* is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. *Cell*, **101**, 543–553.

Elbashir,S.M., Harborth,J., Lendeckel,W., Yalcin,A., Weber,K. and Tuschl,T. (2001a) Duplexes of 21-nucleotide RNAs mediate RNA interference in mammalian cell culture. *Nature*, **411**, 494–498.

Elbashir,S.M., Lendeckel,W. and Tuschl,T. (2001b) RNA interference is mediated by 21 and 22 nt RNAs. *Genes Dev.*, **15**, 188–200.

Fire,A. (1999) RNA-triggered gene silencing. *Trends Genet.*, **15**, 358–363.

Fire,A., Xu,S., Montgomery,M.K., Kostas,S.A., Driver,S.E. and Mello,C.C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, **391**, 806–811.

Fraser,A.G., Kamath,R.S., Zipperlen,P., Martinez-Campos,M., Sohrmann,M. and Ahringer,J. (2000) Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference. *Nature*, **408**, 325–330.

Gönczy,P. *et al.* (2000) Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. *Nature*, **408**, 331–336.

Grishok,A. *et al.* (2001) Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell*, **106**, 23–34.

Hamilton,A.J. and Baulcombe,D.C. (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science*, **286**, 950–952.

Hammond,S.M., Bernstein,E., Beach,D. and Hannon,G.J. (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature*, **404**, 293–296.

Hammond,S.M., Boettcher,S., Caudy,A.A., Kobayashi,R. and Hannon,G.J. (2001a) Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science*, **293**, 1146–1150.

Hammond,S.M., Caudy,A.A. and Hannon,G.J. (2001b) Post-transcriptional gene silencing by double-stranded RNA. *Nature Rev. Genet.*, **2**, 110–1119.

Hutvágner,G., Mlynarova,L. and Nap,J.P. (2000) Detailed characterization of the posttranscriptional gene-silencing-related small RNA in a GUS gene-silenced tobacco. *RNA*, **6**, 1445–1454.

Hutvágner,G., McLachlan,J., Bálint,É., Tuschl,T. and Zamore,P.D. (2001) A cellular function for the RNA interference enzyme dicer in small temporal RNA maturation. *Science*, **293**, 834–838.

Jacobsen,S.E., Running,M.P. and Meyerowitz,M.E. (1999) Disruption of an RNA helicase/RNase III gene in *Arabidopsis* causes unregulated cell division in floral meristems. *Development*, **126**, 5231–5243.

Knight,S.W. and Bass,B.L. (2001) A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *C. elegans*. *Science*, **293**, 2269–2271.

Lam,G. and Thummel,C.S. (2000) Inducible expression of double-stranded RNA directs specific genetic interference in *Drosophila*. *Curr. Biol.*, **10**, 957–963.

Maeda,I., Kohara,Y., Yamamoto,M. and Sugimoto,A. (2001) Large-scale analysis of gene function in *Caenorhabditis elegans* by high-throughput RNAi. *Curr. Biol.*, **11**, 171–176.

Matsuda,S., Ichigotani,Y., Okuda,T., Irimura,T., Nakatsugawa,S. and Hamaguchi,M. (2000) Molecular cloning and characterization of a novel human gene (HERNA) which encodes a putative RNA-helicase. *Biochim. Biophys. Acta*, **1490**, 163–169.

Moss,E.G. (2000) Non-coding RNAs: lightning strikes twice. *Curr. Biol.*, **10**, R436–R439.

Mourrain,P. *et al.* (2000) *Arabidopsis SGS2* and *SGS3* genes are required for posttranscriptional gene silencing and natural virus resistance. *Cell*, **101**, 533–542.

Nykänen,A., Haley,B. and Zamore,P.D. (2001) Nucleotide cofactor requirements and small interfering RNA conformation in the RNA interference pathway. *Cell*, **107**, 309–321.

Olsen,P.H. and Ambros,V. (1999) The lin-4 regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation. *Dev. Biol.*, **216**, 671–680.

Palauqui,J.C., Elmayan,T., Pollien,J.M. and Vaucheret,H. (1997) Systemic acquired silencing: transgene-specific post-transcriptional silencing is transmitted by grafting from silenced stocks to non-silenced scions. *EMBO J.*, **16**, 4738–4745.

Parrish,S., Fleenor,J., Xu,S., Mello,C. and Fire,A. (2000) Functional anatomy of a dsRNA trigger: differential requirement for the two trigger strands in RNA interference. *Mol. Cell*, **6**, 1077–1087.

Pasquinelli,A.E. *et al.* (2000) Conservation of the sequence of *let-7* heterochronic regulatory RNA. *Nature*, **408**, 86–89.

Piano,F., Schetterdagger,A.J., Mangone,M., Stein,L. and Kemphues,K.J. (2000) RNAi analysis of genes expressed in the ovary of *Caenorhabditis elegans*. *Curr. Biol.*, **10**, 1619–1622.

Ray,A., Lang,J.D., Golden,T. and Ray,S. (1996) Short integument (*SIN1*), a gene required for ovule development in *Arabidopsis*, also controls flowering time. *Development*, **122**, 2631–2638.

Reinhart,B.J., Slack,F.J., Basson,M., Pasquinelli,A.E., Bettinger,J.C., Rougvie,A.E., Horvitz,H.R. and Ruvkun,G. (2000) The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature*, **403**, 901–906.

Sharp,P.A. (2001) RNA interference 2001. *Genes Dev.*, **15**, 485–490.

Smardon,A., Spoerke,J., Stacey,S., Klein,M., Mackin,N. and Maine,E. (2000) EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in *C. elegans*. *Curr. Biol.*, **10**, 169–178.

Stark,G.R., Kerr,I.M., Williams,B.R., Silverman,R.H. and Schreiber,R.D. (1998) How cells respond to interferons. *Annu. Rev. Biochem.*, **67**, 227–264.

Tuschl,T. (2001) RNA interference and small interfering RNAs. *ChemBioChem*, **2**, 239–245.

Tuschl,T., Zamore,P.D., Lehmann,R., Bartel,D.P. and Sharp,P.A. (1999) Targeted mRNA degradation by double-stranded RNA *in vitro*. *Genes Dev.*, **13**, 3191–3197.

Ui-Tei,K., Zenno,S., Miyata,Y. and Saigo,K. (2000) Sensitive assay of

RNA interference in *Drosophila* and Chinese hamster cultured cells using firefly luciferase gene as target. *FEBS Lett.*, **479**, 79–82.

Voinnet,O. (2001) RNA silencing as a plant immune system against viruses. *Trends Genet.*, **17**, 449–459.

Voinnet,O., Vain,P., Angell,S. and Baulcombe,D.C. (1998) Systemic spread of sequence-specific transgene RNA degradation in plants is initiated by localized introduction of ectopic promoterless DNA. *Cell*, **95**, 177–187.

Waterhouse,P.M., Wang,M.B. and Lough,T. (2001) Gene silencing as an adaptive defence against viruses. *Nature*, **411**, 834–842.

Wightman,B., Ha,I. and Ruvkun,G. (1993) Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. *Cell*, **75**, 855–862.

Yang,D., Lu,H. and Erickson,J.W. (2000) Evidence that processed small dsRNAs may mediate sequence-specific mRNA degradation during RNAi in *Drosophila* embryos. *Curr. Biol.*, **10**, 1191–1200.

Zamore,P.D., Tuschl,T., Sharp,P.A. and Bartel,D.P. (2000) RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell*, **101**, 25–33.

Received August 23, 2001; revised October 10, 2001; accepted October 16, 2001

EXHIBIT 4

ogy), respectively. Staining specificity was controlled by single staining, as well as by using secondary antibodies in the absence of the primary stain.

Generation of target cells

Target cells displaying a membrane-integral version of either wild-type HEL or a mutant[10] exhibiting reduced affinity for HyHEL10 ([R^{21}, D^{101}, G^{102}, N^{103}] designated HEL*) were generated by transfecting mouse J558L plasmacytoma cells with constructs analogous to those used[10] for expression of soluble HEL/HEL*, except that 14 Ser/Gly codons, the $H2K^b$ transmembrane region, and a 23-codon cytoplasmic domain were inserted immediately upstream of the termination codon by polymerase chain reaction. For mHEL–GFP, we included the EGFP coding domain in the Ser/Gly linker. Abundance of surface HEL was monitored by flow cytometry and radiolabelled-antibody binding using HyHEL5 and D1.3 HEL-specific monoclonal antibodies, for which the mutant HELs used in this work show unaltered affinities[10].

Interaction assays

For B-cell/target interaction assays, splenic B cells from 3-83 or MD4 transgenic mice[28,29] carrying (IgM + IgD) BCRs specific for HEL or $H2K^k$/$H2K^b$were freshly purified on Lympholyte and incubated with a twofold excess of target cells in RPMI, 50 mM HEPES pH 7.4, for the appropriate time at 37 °C before being applied to polylysine-coated slides. Cells were fixed in 4% paraformaldehyde/PBS or methanol and permeabilized with PBS/0.1% Triton X-100 before immunofluorescence. We acquired confocal images using a Nikon E800 microscope attached to BioRad Radiance Plus scanning system equipped with 488-nm and 543-nm lasers, as well as differential interference contrast for transmitted light. GFP fluorescence in living cells in real time was visualized using a Radiance 2000 and Nikon E300 inverted microscope. Images were processed using BioRad Lasersharp 1024 or 2000 software to provide single plane images, confocal projections or slicing.

Antigen presentation

Presentation of HEL epitopes to T-cell hybridomas 2G7 (specific for $I\text{-}E^k$[HEL^{1-16}]) and 1E5 (specific for $I\text{-}E^d$[$HEL^{10A-116}$]) by transfectants of the LK35.2 B-cell hybridoma expressing an HEL-specific IgM BCR was monitored as described[10].

Received 12 December 2000; accepted 30 March 2001.

1. Lanzavecchia, A. Antigen-specific interaction between T and B cells. *Nature* 314, 537–539 (1985).
2. Klaus, G. G., Humphrey, J. H., Kunkl, A. & Dongworth, D. W. The follicular dendritic cell: its role in antigen presentation in the generation of immunological memory. *Immunol. Rev.* 53, 3–28 (1980).
3. Tew, J. G., Kosco, M. H., Burton, G. F. & Szakal, A. K. Follicular dendritic cells as accessory cells. *Immunol. Rev.* 117, 185–211 (1990).
4. Kosco-Vilbois, M. H., Gray, D., Scheidegger, D. & Julius, M. Follicular dendritic cells help resting B cells to become effective antigen-presenting cells: induction of B7/BB1 and upregulation of major histocompatibility complex class II molecules. *J. Exp. Med.* 178, 2055–2066 (1993).
5. Schamel, W. W. & Reth, M. Monomeric and oligomeric complexes of the B cell antigen receptor. *Immunity* 13, 5–14 (2000).
6. Taylor, R. B., Duffus, W. P. H., Raff, M. C. & de Petris, S. Redistribution and pinocytosis of lymphocyte surface immunoglobulin molecules induced by anti-immunoglobulin antibody. *Nature* 233, 225–227 (1971).
7. Schreiner, G. F. & Unanue, E. R. Capping and the lymphocyte: models for membrane reorganisation. *J. Immunol.* 119, 1549–1551 (1977).
8. Cheng, P. C., Dykstra, M. L., Mitchell, R. N. & Pierce, S. K. A role for lipid rafts in B cell antigen receptor signaling and antigen targeting. *J. Exp. Med.* 190, 1549–1560 (1999).
9. Weintraub, B. C. *et al.* Entry of B cell receptor into signaling domains is inhibited in tolerant B cells. *J. Exp. Med.* 191, 1443–1448 (2000).
10. Batista, F. D. & Neuberger, M. S. Affinity dependence of the B cell response to antigen: a threshold, a ceiling, and the importance of off-rate. *Immunity* 8, 751–759 (1998).
11. Nemazee, D. & Burki, K. Clonal deletion of B lymphocytes in a transgenic mouse bearing anti-MHC class I antibody genes. *Nature* 337, 562–566 (1989).
12. Hartley, S. B. *et al.* Elimination from peripheral lymphoid tissues of self-reactive B lymphocytes recognizing membrane-bound antigens. *Nature* 353, 765–769 (1991).
13. Dustin, M. L. *et al.* Low affinity interaction of human and rat T cell adhesion molecule CD2 with its ligands aligns adhering membranes to achieve high physiological affinity. *J. Biol. Chem.* 272, 30889–30898 (1997).
14. Lang, J. *et al.* B cells are exquisitely sensitive to central tolerance and receptor editing by ultralow affinity, membrane-bound antigen. *J. Exp. Med.* 184, 1685–1697 (1996).
15. Valitutti, S., Muller, S., Cella, M., Padovan, E. & Lanzavecchia, A. Serial triggering of many T-cell receptors by a few peptide–MHC complexes. *Nature* 375, 148–151 (1995).
16. Monks, C. R., Freiberg, B. A., Kupfer, H., Sciaky, N. & Kupfer, A. Three-dimensional segregation of supramolecular activation clusters in T cells. *Nature* 395, 82–86 (1998).
17. Wulfing, C. & Davis, M. M. A receptor/cytoskeletal movement triggered by costimulation during T cell activation. *Science* 282, 2266–2269 (1998).
18. Grakoui, A. *et al.* The immunological synapse: a molecular machine controlling T cell activation. *Science* 285, 221–227 (1999).
19. Leupin, O., Zaru, R., Laroche, T., Muller, S. & Valitutti, S. Exclusion of CD45 from the T-cell receptor signaling area in antigen-stimulated T lymphocytes. *Curr. Biol.* 10, 277–280 (2000).
20. Cagan, R. L., Kramer, H., Hart, A. C. & Zipursky, S. L. The bride of sevenless and sevenless interaction: internalization of a transmembrane ligand. *Cell* 69, 393–399 (1992).
21. Huang, J. F. *et al.* TCR-mediated internalization of peptide–MHC complexes acquired by T cells. *Science* 286, 952–954 (1999).
22. Hwang, I. *et al.* T cells can use either T cell receptor or CD28 receptors to absorb and internalize cell surface molecules derived from antigen-presenting cells. *J. Exp. Med.* 191, 1137–1148 (2000).
23. Batista, F. D. & Neuberger, M. S. B cells extract and present immobilized antigen: implications for affinity discrimination. *EMBO J.* 19, 513–520 (2000).
24. Casten, L. A., Lakey, E. K., Jelachich, M. L., Margoliash, E. & Pierce, S. K. Anti-immunoglobulin augments the B-cell antigen-presentation function independently of internalization of receptor-antigen complex. *Proc. Natl Acad. Sci. USA* 82, 5890–5894 (1985).
25. Siemasko, K., Eisfelder, B. J., Williamson, E., Kabak, S. & Clark, M. R. Signals from the B lymphocyte antigen receptor regulate MHC class II containing late endosomes. *J. Immunol.* 160, 5203–5208 (1998).
26. Serre, K. *et al.* Efficient presentation of multivalent antigens targeted to various cell surface molecules of dendritic cells and surface Ig of antigen-specific B cells. *J. Immunol.* 161, 6059–6067 (1998).
27. Green, S. M., Lowe, A. D., Parrington, J. & Karn, J. Transformation of growth factor-dependent myeloid stem cells with retroviral vectors carrying c-myc. *Oncogene* 737–751 (1989).
28. Russell, D. M. *et al.* Peripheral deletion of self-reactive B cells. *Nature* 354, 308–311 (1991).
29. Goodnow, C. C. *et al.* Altered immunoglobulin expression and functional silencing of self-reactive B lymphocytes in transgenic mice. *Nature* 334, 676–682 (1988).
30. Aluvihare, V. R., Khamlichi, A. A., Williams, G. T., Adorini, L. & Neuberger, M. S. Acceleration of intracellular targeting of antigen by the B-cell antigen receptor: importance depends on the nature of the antigen-antibody interaction. *EMBO J.* 16, 3553–3562 (1997).

Supplementary information is available on *Nature*'s World-Wide Web site (http://www.nature.com) or as paper copy from the London editorial office of *Nature*.

Acknowledgements

We thank B. Amos and S. Reichelt for help and advice with confocal microscopy, and S. Munro for helpful discussions. We are indebted those who provided antibodies, transgenic mice and cell lines. F.D.B. and D.I. were supported by the Arthritis Research Campaign and Studienstiftung des deutschen Volkes, respectively.

Correspondence and requests for materials should be addressed to F.D.B. (e-mail:fdb@mrc-lmb.cam.ac.uk) or M.S.N. (e-mail:msn@mrc-lmb.cam.ac.uk)

Duplexes of 21–nucleotide RNAs mediate RNA interference in cultured mammalian cells

Sayda M. Elbashir*, Jens Harborth†, Winfried Lendeckel*, Abdullah Yalcin*, Klaus Weber† & Thomas Tuschl*

** Department of Cellular Biochemistry; and † Department of Biochemistry and Cell Biology; Max-Planck-Institute for Biophysical Chemistry, Am Fassberg 11, D-37077 Göttingen, Germany*

RNA interference (RNAi) is the process of sequence-specific, post-transcriptional gene silencing in animals and plants, initiated by double-stranded RNA (dsRNA) that is homologous in sequence to the silenced gene[1–4]. The mediators of sequence-specific messenger RNA degradation are 21- and 22-nucleotide small interfering RNAs (siRNAs) generated by ribonuclease III cleavage from longer dsRNAs[5–9]. Here we show that 21-nucleotide siRNA duplexes specifically suppress expression of endogenous and heterologous genes in different mammalian cell lines, including human embryonic kidney (293) and HeLa cells. Therefore, 21-nucleotide siRNA duplexes provide a new tool for studying gene function in mammalian cells and may eventually be used as gene-specific therapeutics.

Uptake of dsRNA by insect cell lines has previously been shown to 'knock-down' the expression of specific proteins, owing to sequence-specific, dsRNA-mediated mRNA degradation[6,10–12]. However, it has not been possible to detect potent and specific RNA interference in commonly used mammalian cell culture systems, including 293 (human embryonic kidney), NIH/3T3 (mouse fibroblast), BHK-21 (Syrian baby hamster kidney), and CHO-K1 (Chinese hamster ovary) cells, applying dsRNA that varies in size between 38 and 1,662 base pairs (bp)[10,12]. This apparent lack of RNAi in mammalian cell culture was unexpected, because RNAi exists in mouse oocytes and early embryos[13,14], and because RNAi-related, transgene-mediated co-suppression was also observed in cultured Rat-1 fibroblasts[15]. But it is known that dsRNA in the cytoplasm of mammalian cells can trigger profound physiological

uridine residues. The thymidine overhang was chosen because it reduces costs of RNA synthesis and may enhance nuclease resistance of siRNAs in the cell culture medium and within transfected cells. As in the *Drosophila in vitro* system (data not shown), substitution of uridine by thymidine in the 3′ overhang was well tolerated in cultured mammalian cells (Fig. 2a, c, e, g and i), and the sequence of the overhang appears not to contribute to target recognition[9].

In co-transfection experiments, 25 nM siRNA duplexes were used (Figs 2 and 3; concentration is in respect to the final volume of tissue culture medium). Increasing the siRNA concentration to 100 nM did not enhance the specific silencing effects, but started to affect transfection efficiencies, perhaps due to competition for liposome encapsulation between plasmid DNA and siRNA (data not shown). Decreasing the siRNA concentration to 1.5 nM did not reduce the specific silencing effect (data not shown), even though the siRNAs were now only 2- to 20-fold more concentrated than the DNA plasmids; the silencing effect only vanishes completely if the siRNA concentration was dropped below 0.05 nM. This indicates that siRNAs are extraordinarily powerful reagents for mediating gene silencing, and that siRNAs are effective at concentrations that are several orders of magnitude below the concentrations applied in conventional antisense or ribozyme gene-targeting experiments[20].

To monitor the effect of longer dsRNAs on mammalian cells, 50- and 500-bp dsRNAs that are cognate to the reporter genes were prepared. As a control for nonspecific inhibition, dsRNAs from humanized GFP (hG)[21] was used. In these experiments, the reporter plasmids were co-transfected with either 0.21 μg siRNA duplexes or 0.21 μg longer dsRNAs. The siRNA duplexes only reduced the expression of their cognate reporter gene, while the longer dsRNAs strongly and nonspecifically reduced reporter-gene expression. The effects are illustrated for HeLa S3 cells as a representative example (Fig. 3a and b). The absolute luciferase activities were decreased nonspecifically 10- to 20-fold by 50-bp dsRNA, and 20- to 200-fold by 500-bp dsRNA co-transfection, respectively. Similar nonspecific effects were observed for COS-7 and NIH/3T3 cells. For 293 cells, a 10- to 20-fold nonspecific reduction was observed only for 500-bp dsRNAs. Nonspecific reduction in reporter-gene expression by dsRNA > 30 bp was expected as part of the interferon response[16]. Interestingly, superimposed on the nonspecific interferon response, we detect additional sequence-specific, dsRNA-mediated silencing. The sequence-specific silencing effect of long dsRNAs, however, became apparent only when the relative reporter-gene activities were normalized to the hG dsRNA controls (Fig. 3c). Sequence-specific silencing by 50- or 500-bp dsRNAs reduced the targeted reporter-gene expression by an additional 2- to 5-fold. Similar effects were also detected in the other three mammalian cell lines tested (data not shown). Specific silencing effects with dsRNAs (356–1,662 bp) were previously reported in CHO-K1 cells, but the amounts of dsRNA required to detect a 2- to 4-fold specific reduction were about 20-fold higher than in our experiments[12]. Also, CHO-K1 cells appear to be deficient in the interferon response. In another report, 293, NIH/3T3 and BHK-21 cells were tested for RNAi using luciferase/β-galactosidase (lacZ) reporter combinations and 829-bp specific lacZ or 717-bp nonspecific green fluorescent protein (GFP) dsRNA[10]. The lack of detected RNAi in this case may be due to the less sensitive luciferase/lacZ reporter assay and the length differences of target and control dsRNA. Taken together, our results indicate that RNAi is active in mammalian cells, but that the silencing effect is difficult to detect if the interferon system is activated by dsRNA > 30 bp.

To test for silencing of endogenous genes, we chose four genes coding for cytoskeletal proteins: lamin A/C, lamin B1, nuclear mitotic apparatus protein (NuMA) and vimentin[27]. The selection was based on the availability of antibodies needed to quantitate the silencing effect. Silencing was monitored 40 to 45 h after transfection to allow for turnover of the protein of the targeted genes. As shown in Fig. 4, the expression of lamin A/C was specifically reduced by the cognate siRNA duplex (Fig. 4a), but not when nonspecific siRNA directed against firefly luciferase (Fig. 4b) or buffer (Fig. 4c) was used. The expression of a non-targeted gene, NuMA, was unaffected in all treated cells (Fig. 4d–f), demonstrating the integrity of the targeted cells. The reduction in lamin A/C proteins was more than 90% complete as quantified by western blotting (Fig. 4j, k). We note that lamin A/C 'knock-out' mice are

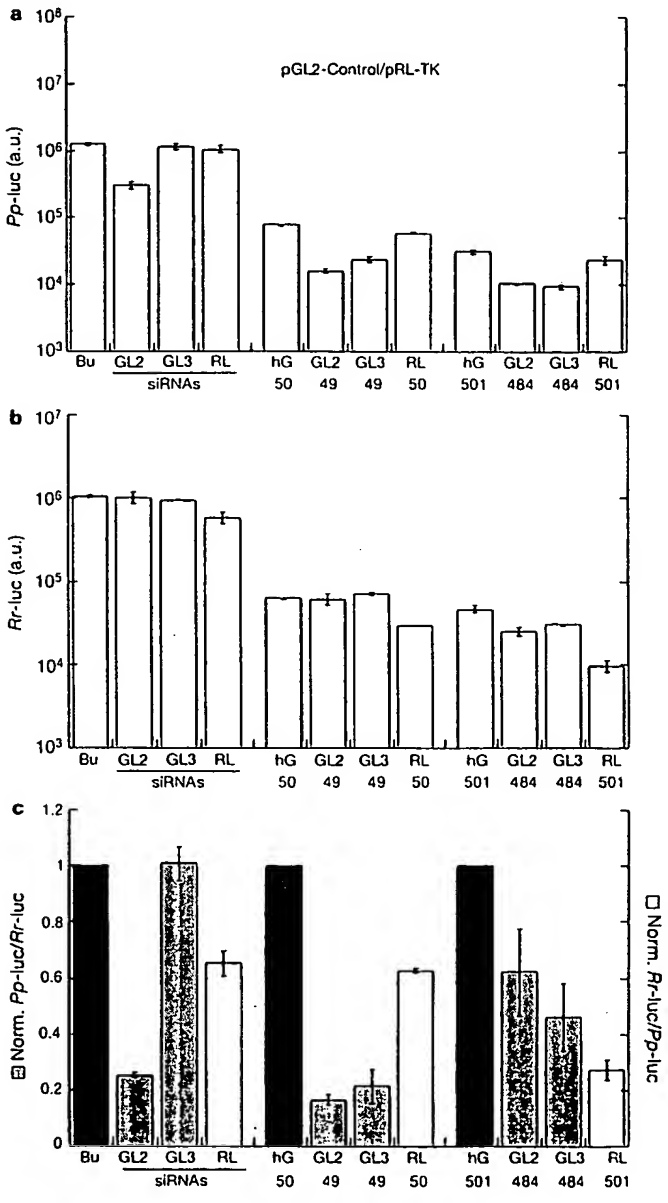


Figure 3 Effects of 21-nucleotide siRNAs, 50-bp, and 500-bp dsRNAs on luciferase expression in HeLa cells. The exact length of the long dsRNAs in base pairs is indicated below the bars. Experiments were performed with pGL2-Control and pRL-TK reporter plasmids. The data were averaged from two independent experiments ± s.d. **a**, Absolute *Pp*-luc expression, plotted in arbitrary luminescence units (a.u.). **b**, *Rr*-luc expression, plotted in arbitrary luminescence units. **c**, Ratios of normalized target to control luciferase. The ratios of luciferase activity for siRNA duplexes were normalized to a buffer control (Bu, black bars); the luminescence ratios for 50- or 500-bp dsRNAs were normalized to the respective ratios observed for 50- and 500-bp dsRNAs from humanized GFP (hG, black bars). We note that the overall differences in sequence between the 49- and 484-bp GL2 and GL3 dsRNAs are not sufficient to confer specificity for targeting GL2 and GL3 targets (43-nucleotide uninterrupted identity in 49-bp segment, 239-nucleotide longest uninterrupted identity in 484-bp segment)[30].

viable for a few weeks after birth[23] and that the lamin A/C knockdown in cultured cells was not expected to cause cell death. Lamin A and C are produced by alternative splicing in the 3′ region and are present in equal amounts in the lamina of mammalian cells (Fig. 4j, k). Transfection of siRNA duplexes targeting lamin B1 and NuMA reduced the expression of these proteins to low levels (data not shown), but we were not able to observe a reduction in vimentin expression. This could be due to the high abundance of vimentin in the cells (several per cent of total cell mass) or because the siRNA sequence chosen was not optimal for targeting of vimentin.

The mechanism of the 21-nucleotide siRNA-mediated interference process in mammalian cells remains to be uncovered, and silencing might occur post-transcriptionally and/or transcriptionally. In *Drosophila* lysate, siRNA duplexes mediate post-transcriptional gene silencing by reconstitution of siRNA-protein complexes (siRNPs), which guide mRNA recognition and targeted cleavage[6,7,9]. In plants, dsRNA-mediated post-transcriptional silencing has also been linked to DNA methylation, which may also be directed by 21-nucleotide siRNAs[24]. Methylation of promoter regions can lead to transcriptional silencing[25], but methylation in coding sequences does not[26]. DNA methylation and transcriptional silencing in mammals are well documented processes[27], yet their mechanisms have not been linked to that of post-transcriptional silencing. Methylation in mammals is predominantly directed towards CpG dinucleotide sequences. There is no CpG sequence in the RL or lamin A/C siRNA, although both siRNAs mediate specific silencing in mammalian cell culture, so it is unlikely that DNA methylation is essential for the silencing process.

Thus we have shown, for the first time, siRNA-mediated gene silencing in mammalian cells. The use of exogenous 21-nucleotide siRNAs holds great promise for analysis of gene function in human cell culture and the development of gene-specific therapeutics. It will also be of interest in understanding the potential role of endogenous siRNAs in the regulation of mammalian gene function. □

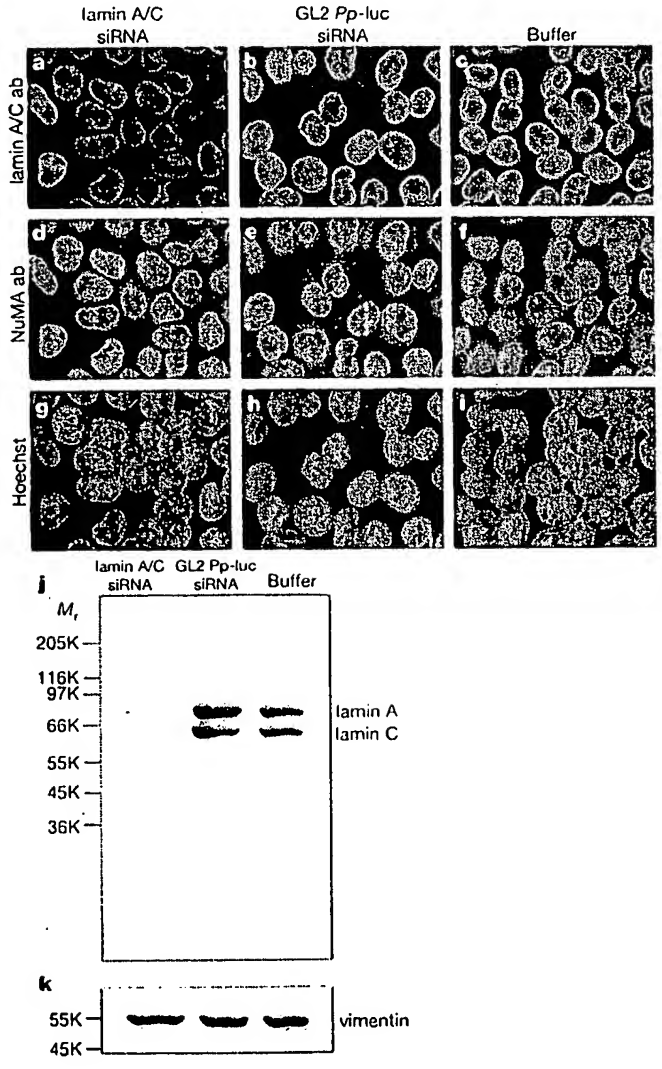


Figure 4 Silencing of nuclear envelope proteins lamin A/C in HeLa cells. Triple fluorescence staining of cells transfected with lamin A/C siRNA duplex (**a**, **d**, **g**), with GL2 luciferase siRNA duplex (nonspecific siRNA control) (**b**, **e**, **h**), and with buffer only (**c**, **f**, **i**). **a**–**c**, Staining with lamin A/C specific antibody; **d**–**f**, staining with NuMA-specific antibody; **g**–**i**, Hoechst staining of nuclear chromatin. Bright fluorescent nuclei in a represent untransfected cells. **j**, **k**, Western blots of transfected cells using lamin A/C- (**j**) or vimentin-specific (**k**) antibodies. The Western blot was stripped and re-probed with vimentin antibody to check for equal loading of total protein.

Methods

RNA preparation

21-nucleotide RNAs were chemically synthesized using Expedite RNA phosphoramidites and thymidine phosphoramidite (Proligo, Germany). Synthetic oligonucleotides were deprotected and gel-purified". The accession numbers given below are from GenBank. The siRNA sequences targeting GL2 (Acc. No. X65324) and GL3 luciferase (Acc. No. U47296) corresponded to the coding regions 153–173 relative to the first nucleotide of the start codon; siRNAs targeting RL (Acc. No. AF025846) corresponded to region 119–139 after the start codon. The siRNA sequence targeting lamin A/C (Acc. No. X03444) was from position 608–630 relative to the start codon; lamin B1 (Acc. No. NM_005573) siRNA was from position 672–694; NuMA (Acc. No. Z11583) siRNA from position 3,988–4,010, and vimentin (Acc. No. NM_003380) from position 346–368 relative to the start codon. Longer RNAs were transcribed with T7 RNA polymerase from polymerase chain reaction (PCR) products, followed by gel purification. The 49- and 484-bp GL2 or GL3 dsRNAs corresponded to positions 113–161 and 113–596, respectively, relative to the start of translation; the 50- and 501-bp RL dsRNAs corresponded to position 118–167 and 118–618, respectively. PCR templates for dsRNA synthesis targeting humanized GFP (hG) were amplified from pAD3 (ref. 21), whereby 50- and 501-bp hG dsRNA corresponded to positions 121–170 and 121–621, respectively, to the start codon.

For annealing of siRNAs, 20 μM single strands were incubated in annealing buffer (100 mM potassium acetate, 30 mM HEPES-KOH at pH 7.4, 2 mM magnesium acetate) for 1 min at 90 °C followed by 1 h at 37 °C. The 37 °C incubation step was extended overnight for the 50- and 500-bp dsRNAs, and these annealing reactions were performed at 8.4 μM and 0.84 μM strand concentrations, respectively.

Cell culture

S2 cells were propagated in Schneider's *Drosophila* medium (Life Technologies) supplemented with 10% fetal bovine serum (FBS) 100 units ml^{-1} penicillin, and 100 μg ml^{-1} streptomycin at 25 °C. 293, NIH/3T3, HeLa S3, HeLa SS6, COS-7 cells were grown at 37 °C in Dulbecco's modified Eagle's medium supplemented with 10% FBS, 100 units ml^{-1} penicillin, and 100 μg ml^{-1} streptomycin. Cells were regularly passaged to maintain exponential growth. Twenty-four h before transfection at 50–80% confluency, mammalian cells were trypsinized and diluted 1:5 with fresh medium without antibiotics ($1–3 \times 10^5$ cells ml^{-1}) and transferred to 24-well plates (500 μl per well). S2 cells were not trypsinized before splitting. Co-transfection of reporter plasmids and siRNAs was carried out with Lipofectamine 2000 (Life Technologies) as described by the manufacturer for adherent cell lines. Per well, 1.0 μg pGL2-Control (Promega) or pGL3-Control (Promega), 0.1 μg pRL-TK (Promega), and 0.21 μg siRNA duplex or dsRNA, formulated into liposomes, were applied; the final volume was 600 μl per well. Cells were incubated 20 h after transfection and appeared healthy thereafter. Luciferase expression was subsequently monitored with the Dual luciferase assay (Promega). Transfection efficiencies were determined by fluorescence microscopy for mammalian cells lines after co-transfection of 1.1 μg hGFP-encoding pAD3 (ref. 21) and 0.21 μg inverted GL2 siRNA, and were 70–90%. Reporter plasmids were amplified in XL-1 Blue (Stratagene) and purified using the Qiagen EndoFree Maxi Plasmid Kit.

Transfection of siRNAs for targeting endogenous genes was carried out using Oligofectamine (Life Technologies) and 0.84 μg siRNA duplex per well, but it was recently found that as little as 0.01 μg siRNAs per well are sufficient to mediate silencing. HeLa SS6 cells were transfected one to three times in approximately 15 h intervals and were assayed 40 to 45 h after the first transfection. It appears, however, that a single transfection is as efficient as multiple transfections. Transfection efficiencies as determined by immunofluorescence of targeted cells were in the range of 90%. Specific silencing of targeted genes was confirmed by at least three independent experiments.

Western blotting and immunofluorescence microscopy

Monoclonal 636 lamin A/C specific antibody[28] was used as undiluted hybridoma supernatant for immunofluorescence and 1/100 dilution for western blotting. Affinity-purified polyclonal NuMA protein 705 antibody[29] was used at a concentration of 10 μg ml^{-1} for

EXHIBIT 5

MOLECULAR AND CELLULAR BIOLOGY, Apr. 1995, p. 1907–1914
0270-7306/95/$04.00+0

Vol. 15, No. 4

Tetracycline-Reversible Silencing of Eukaryotic Promoters

ULRICH DEUSCHLE,* WOLFRAM K.-H. MEYER, AND HANS-JÜRGEN THIESEN

Basel Institute for Immunology, CH-4005 Basel, Switzerland

Received 12 September 1994/Returned for modification 10 November 1994/Accepted 4 January 1995

A tetracycline-controlled transrepressor protein has been engineered to silence transcriptional activities of eukaryotic promoters that are stably integrated into the chromatin of human cells. By fusing the KRAB domain of human Kox1 to the Tet repressor derived from Tn*10* of *Escherichia coli*, a tetracycline-controlled hybrid protein (TetR-KRAB) was generated and constitutively expressed in HeLa cells. The TetR-KRAB protein binds to *tet* operator (*tetO*) sequences in the absence but not in the presence of tetracycline. When TetR-KRAB bound to *tetO* sequences upstream of the immediate-early promoter-enhancer of human cytomegalovirus (CMV), the expression of a CMV-driven luciferase reporter construct (ptetO7-CMV-L) was repressed in transient transfection experiments. This silencing was found to operate on different promoters and from *tetO* sequences placed more than 3 kb from the transcriptional start site. We constructed a stable, doubly transfected cell line (TIS-10) carrying a chromosomally integrated ptetO7-CMV-L reporter construct and expressing the TetR-KRAB protein. Upon addition of tetracycline, luciferase expression was induced more than 50-fold above the baseline level, with half-maximal induction by 2 days. Furthermore, a protein of around 110 kDa was found to coimmunoprecipitate with the TetR-KRAB fusion protein. This protein might play a role as an adaptor protein mediating the silencing exerted by the TetR-KRAB protein. The TetR-KRAB silencing system should be useful as a genetic switch for regulating the expression of chromosomally integrated heterologous and endogenous genes present in mammalian genomes.

Inducible gene expression has been a valuable tool for the study of gene function in bacteria, yeasts, and *Drosophila melanogaster*. In mammalian cells, eukaryotic promoter systems that respond to inducing agents such as glucocorticoid hormone (19, 23), heat shock (40), heavy metal ions (28), or interferon (32) have been used. To circumvent limitations due either to the leakiness of utilized promoters or the pleiotropic effects of inducing agents, chimeric transcription factors have been generated, for example, by fusing the strong transcriptional activator domain of VP16 to the Tn*10*-derived prokaryotic tetracycline repressor (TetR) protein (16). In general, chimeric transcription factors of this kind, TetR-VP16 (16), GAL4-VP16 (33), LexA-VP16 (7), and LacR-VP16 (22), are targeted to minimal promoters which are fused to multiple DNA binding sites specific for the factor in question. This results in marked transcriptional stimulation of the respective minimal promoter. In order to modulate the transcriptional activity of chimeric transcription factors, movable steroid binding domains originating from the glucocorticoid receptor (30) or the estrogen receptor (38) have been fused to sequence-specific DNA binding domains. In the case of a GAL4-VP16-estrogen receptor fusion, estrogen administration has been shown to produce a functionally active chimeric transcription factor that activates a target promoter (6).

In the tetracycline system, the TetR protein fused to the transactivating domain of VP16 (called tTA) has been shown to bind to and strongly activate minimal promoter systems containing seven tetracycline operator (*tetO*) sequences (15, 16). The binding of tTA to the *tetO* sequences is blocked by tetracycline, preventing the activation of the target promoter which generates a very tight genetic switch (16).

In this report, we present a novel system for tetracycline-controlled silencing of eukaryotic promoters. Here, the KRAB repressor domain of the human Kox1 zinc finger protein (27) has been fused to the TetR protein. In the absence of tetracycline, this chimeric DNA-binding protein (TetR-KRAB) exerts its silencing activity by binding to several *cis*-acting *tetO* sites placed at a distance from the transcriptional initiation site of a eukaryotic promoter. Promoter activity is restored upon administration of tetracycline, which prevents binding of TetR-KRAB to the *tetO* sequences. We also present evidence that the TetR-KRAB forms a complex with a protein of around 110 kDa. This protein, provisionally called SMP1 (silencing-mediating protein 1), might constitute a putative corepressor protein that mediates the repression activity employed by the KRAB domain of Kox1.

MATERIALS AND METHODS

Plasmid constructions. The fusion between TetR and 121 amino acids of the NH_2 terminus of Kox1 was constructed by PCR amplification. A PCR fragment was generated by using two synthetic oligonucleotides, 5'ATCAGGAATTC AACCATGGCTAGATTAGATAAAAG3' and 5'GTCCTGTCGACCTTTCTC TTCTTTTTTGGCGACCCACTTTCACATTT3', and pUHD15-1 (16), introducing an *Nco*I site around the translational start of the TetR preceded by an *Eco*RI restriction site (sites are underlined). In addition, six amino acids comprising the simian virus 40 (SV40) large-T-antigen nuclear localization sequence (20) (nucleotides in boldface) were fused in frame with the last amino acid residue of TetR followed by a *Sal*I restriction site (underlined). The resulting PCR fragment was digested with *Eco*RI and *Sal*I. A second PCR fragment was generated by using two oligonucleotides, 5'GACAGGTCGACGGCGGTGGT GCTTTGTC3' and 5'GCTGCGGATCCTTAAACTGATGATTTGATTTC3', and plasmid pKox1 (35), introducing a *Sal*I site at the NH_2-terminal end of the KRAB domain of Kox1 and a stop codon at position 121 of Kox1 followed by a *Bam*HI restriction site. After digestion with *Sal*I and *Bam*HI, both fragments were ligated into *Eco*RI- and *Bam*HI-digested pUHD10-1 (12), generating pCMV-tetR-KRAB. To generate pCMV-tetR-KRAB-hyg, a *Hin*dIII-*Bsm*I fragment of plasmid p220.3, a derivative of p220.1 (13) (kindly provided by M. Gossen and H. Bujard) containing the hygromycin resistance gene under control of the herpes simplex virus (HSV) thymidine kinase (TK) promoter, was introduced into a *Pvu*II site downstream of the polyadenylation signal in pCMV-tetR-KRAB. To generate pCMV-tetR, pCMV-tetR-KRAB was digested with *Sal*I and *Bam*HI, and the ends were filled in and religated. This retains the fusion of the complete TetR to the nuclear localization signal of SV40 large T antigen (Pro-Lys-Lys-Lys-Arg-Lys) followed by five amino acids originating from the vector backbone (Val-Glu-Ile-Gln-Thr-Stop).

Plasmid ptetO7-CMV-L was generated by ligating the 786-bp *Xho*I-*Bam*HI

* Corresponding author. Present address: Pharmaceutical Research Gene Technologies, F. Hoffmann-La Roche Ltd., CH-4002 Basel, Switzerland. Phone: 41-61-6886874. Fax: 41-61-6881448.

cytomegalovirus (CMV) promoter fragment of pUHD10-1 (12) into *Xho*I-*Bam*HI-digested pUHC13-4 (16). This places seven *tetO* sequences just upstream of the human CMV immediate-early promoter/enhancer (4). Plasmid ptetO7-TK-L was generated by ligating an *Xho*I-*Bgl*II fragment from pT109luc (29) comprising the HSV TK promoter (−109 to +52, both GC boxes and the CCAAT box present) into *Xho*I-*Bam*HI-digested pUHC13-4 (16).

Cell culture and transfections. HeLa cells were grown in Iscove's modified Dulbecco medium (GIBCO) supplemented with 5% fetal calf serum (FCS; Boehringer Mannheim) in a 5% CO_2 environment at 37°C. Cells were plated at a density of 10^5 cells per 35-mm-diameter dish and transfected at 80% confluence with 5 μg of plasmid DNA (4.5 μg of the expression and 0.5 μg of the reporter plasmid) as a calcium phosphate precipitate (1). After the transfection, cells were further incubated in Iscove's modified Dulbecco medium supplemented with 5% FCS and no or 0.5 μg of tetracycline (Sigma) per ml for 24 h before determination of luciferase activities from cell extracts. Determination of β-galactosidase activities (1) from cotransfected pCH110 (2 μg per dish; Pharmacia) was used to control for transfection efficiencies.

To establish stable cell lines expressing the TetR-KRAB fusion protein, HeLa cells were transfected with 2.5 μg of pCMV-tetR-KRAB-hyg DNA linearized with *Asp* 700. Forty-eight hours after transfection, cells were trypsinized and split at a ratio of 1:10, and resistant colonies were selected in Iscove modified Dulbecco medium supplemented with 5% FCS (Boehringer) and 250 μg of hygromycin B (Sigma) per ml. Resistant colonies were subcloned by limiting dilution in nonselective medium and tested for expression of functional TetR-KRAB protein, resulting in the A12 cell line. To establish a cell line containing the ptetO7-CMV-L construct together with pCMV-tetR-KRAB, A12 cells were transfected with 5 μg of a 15:1 mixture of ptetO7-CMV-L linearized with *Hin*dIII and pNEO5′ (26). Colonies resistant to G418 (500 μg/ml) were grown in presence of 0.5 μg of tetracycline per ml, and luciferase activities were determined from extracts. One positive clone, called TIS-10, was chosen for further characterization.

Luciferase assays. Cells were grown in absence or presence of tetracycline (0.5 μg/ml) in 35-mm-diameter dishes, washed with 2 ml of phosphate-buffered saline (PBS), and then lysed in 500 μl of 20 mM Tris-HCl (pH 7.8)–150 mM NaCl–1 mM EDTA–1 mM dithiothreitol–0.6% Nonidet P-40. The lysate was collected and centrifuged in an Eppendorf centrifuge for 10 s. Aliquots (10 μl) of the supernatants were mixed with 350 μl of 25 mM glycylglycine–15 mM $MgSO_4$–5 mM ATP and assayed for luciferase activity (14) in a Lumat LB9501 (Berthold, Wildbad, Germany), using the integral mode (10 s). D-Luciferin (L 6882; Sigma) was used at 1 mM. Protein content of the lysates was determined by the method of Bradford (5).

Generation of Kox1 antiserum. Recombinant Kox1 protein was generated by using the T7 expression vector pAR3039 (31). Briefly, the *Bam*HI cDNA fragment of Kox1 (35) was cloned into the *Bam*HI site of pAR3039, expressed in the *Escherichia coli* host strain BL21(DE3), and the Kox1 protein was purified by Mono S fast protein liquid chromatography (36). Functionally active protein was obtained after renaturing Kox1 in a dialyzing procedure adapted for zinc finger proteins (28a). One rabbit was immunized with renatured Kox1 protein.

Immunofluorescence. HeLa A12 cells were grown in tissue culture chamber slides (Miles Scientific), washed with PBS, fixed in 3% paraformaldehyde in PBS for 10 min, washed three times with PBS, and permeabilized with 0.6% Nonidet P-40 in PBS for 5 min. After washing with PBS, the fixed monolayer was blocked by incubation with 3% bovine serum albumin (BSA) in PBS for 20 min to reduce nonspecific antibody adsorption. Cells were then washed with PBS and incubated for 40 min with the rabbit anti-human Kox1 immune serum diluted 1/500 in 3% BSA in PBS. After four washes with PBS, cells were incubated for 20 min in the dark with a 1/100 dilution of a fluorescein-conjugated goat anti-rabbit immunoglobulin G (Boehringer Mannheim) in 3% PBS. Finally, cells were washed five times with PBS and covered with 4 μl of 90% glycerol–1 mM KI–0.2 μg of 4′,6-diamidino-2-phenylindole (DAPI) per ml (for staining of the DNA). Coverslips were applied and sealed in place with clear nail polish. Cells were visualized by fluorescence microscopy with a Zeiss microscope at a magnification of 1:200.

Electrophoretic mobility shift assay. A12 cells were grown in 35-mm-diameter dishes to 80% confluence, washed with PBS, and scraped off the plate, and the sedimented cells were resuspended in 200 μl of a buffer containing 20 mM *N*-2-hydroxyethylpiperazine-N′-2-ethanesulfonic acid (HEPES)-KOH (pH 7.9), 25% glycerol, 400 mM NaCl, 1 mM EDTA, 1 mM EGTA, 0.1 mM Na_3VaO_4, 25 mM β-glycerophosphate, 1 mM dithiothreitol, and 0.5 mM phenylmethylsulfonyl fluoride. After freeze-thawing and incubation for 10 min on ice, the cellular debris was sedimented (10 min, 13,000 × *g*). Aliquots of total cell extract (2 μg of total protein) were mixed with 2 fmol of ^{32}P-labeled $tetO_2$ DNA (16) in 25 μl of binding buffer containing 10 mM HEPES-KOH (pH 7.9), 1 mM EDTA, 4% Ficoll 400, and 4 μg of sonicated calf thymus DNA. After 15 min at room temperature, the binding reaction mixture was loaded onto a 4% polyacrylamide–0.2% bisacrylamide gel with 0.25× Tris-borate-EDTA (1) as the running buffer. Polyacrylamide gel electrophoresis (PAGE) was performed at room temperature at 7.5 V/cm. For assays in which antibodies were included, binding was allowed for 10 min before addition of 2.5 μl of a 1:10 dilution of serum in binding buffer, and incubation was continued for an additional 5 min at room temperature. When deoxycholate (DOC) was used, appropriate dilutions of the DOC in binding buffer were added after 10 min incubation in absence of DOC. Dried gels were visualized by autoradiography using Kodak AR-5 film.

Immunoprecipitation. HeLa and HeLa A12 cells were grown in 100-mm-diameter dishes to 80% confluence, washed, and grown for 2 h in 4 ml of L-methionine-free modified Eagle's medium (GIBCO) supplemented with 5% dialyzed FCS and 500 μCi of L-[^{35}S]methionine (Amersham). Following the addition of 4 ml of Dulbecco's modified Eagle's medium (GIBCO) supplemented with 5% FCS, cells were grown for an additional 30 min, washed with PBS, harvested, and lysed by freeze-thawing in 500 μl of immunoprecipitation buffer (IP buffer; 50 mM Tris-HCl [pH 7.4], 250 mM NaCl, 0.05% Nonidet P-40, 5 mM EDTA, 50 mM NaF, 1 mM Na_3VaO_4, 0.5 mM phenylmethylsulfonyl fluoride, 2 mM dithiothreitol) followed by centrifugation for 10 min at 10,000 × *g*. The crude supernatant was adjusted to 1% BSA in IP buffer and precleared with 50 μl of a 50% slurry of protein G-agarose beads (Boehringer). The supernatant was divided in two aliquots; to each, 2 μl of the anti-Kox1 rabbit immune serum was added. After 2 h of incubation at 4°C, a slurry of 50 μl of protein G-agarose was added, and the mixture was incubated for an additional 30 min. The beads were collected and washed three times in IP buffer. To one aliquot of the beads, 50 μl of IP buffer containing 0.1% DOC was added; after 5 min at room temperature, the beads were separated from the supernatant. Fifty microliters of IP buffer was added to the beads, and the samples were adjusted to SDS-sample buffer and separated on an SDS–12% polyacrylamide gel (1). The gel was soaked in Amplify (Amersham), and the labeled proteins were visualized by autoradiography.

RESULTS

Construction of expression plasmids. It was recently shown that a fusion protein of the yeast GAL4 DNA binding domain and the KRAB domain of Kox1 expressed in mammalian cells strongly repressed a cotransfected chloramphenicol acetyltransferase reporter construct containing GAL4 binding sites upstream of the TK promoter (27). We decided to take advantage of this finding. We fused the first NH_2-terminal 121 amino acid residues of Kox1, containing the KRAB domain (35), to the C terminus of the prokaryotic TetR encoded by Tn*10* from *E. coli* (16) to create a tetracycline-responsive transcriptional silencer protein, TetR-KRAB. As a linker between the two domains, six amino acids comprising the nuclear localization sequence of SV40 large T antigen (20) were added to favor nuclear localization of the TetR-KRAB fusion protein in mammalian cells. In plasmids used for expression of TetR-KRAB (pCMV-tetR-KRAB and pCMV-tetR-KRAB-hyg) in mammalian cells, the sequence encoding TetR-KRAB is flanked upstream by the human CMV immediate-early promoter and downstream by the SV40 poly(A) site. As a control, the KRAB domain was removed, creating TetR fused to the nuclear localization signal followed by five additional amino acids in pCMV-tetR. The reporter plasmids (ptetO7-CMV-L and ptetO7-TK-L) are based on pUHC13-4 (16). In these constructs, the CMV or HSV TK promoter was inserted downstream of seven *tetO* sequences and upstream of the luciferase as a reporter gene. The right boundary of the fragment containing the seven *tetO* sites is located more than 685 and 109 bp upstream of the transcriptional initiation site of the CMV and TK promoters, respectively. Relevant features of plasmids pCMV-tetR-KRAB, pCMV-tetR, and ptetO7-CMV-L are outlined in Fig. 1. The rationale is to generate a fusion protein TetR-KRAB that binds in *cis* to *tetO* sequences upstream of the transcriptional initiation site of the CMV promoter (Fig. 1) or to any other promoter. Binding of the KRAB domain to the *tetO* sites should result in active repression of productive transcription measured by luciferase activity. Addition of low levels of tetracycline to the growth medium should prevent TetR-KRAB from binding to the *tetO* sequences and should thereby restore full productive transcriptional initiation from the respective promoter.

Repression of transcription mediated by the TetR-KRAB fusion protein. HeLa cells were cotransfected with the reporter plasmid ptetO7-CMV-L and one of the expression plasmids pCMV, pCMV-tetR-KRAB, and pCMV-tetR and grown with

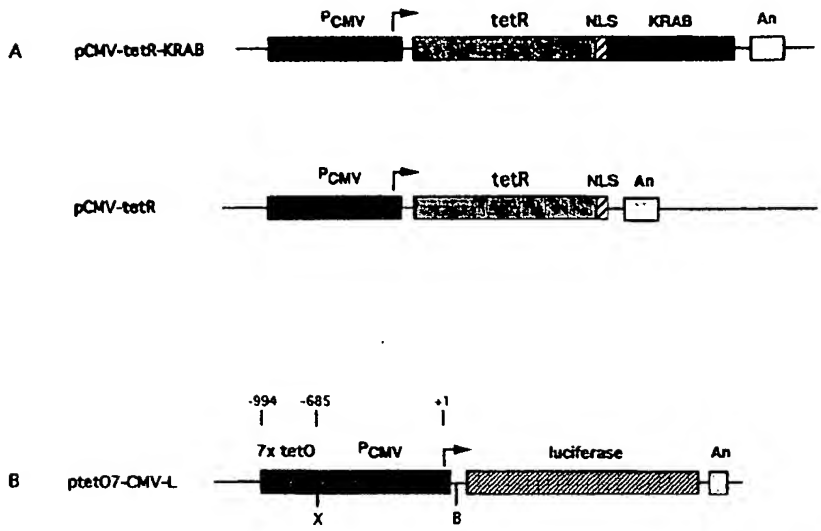


FIG. 1. Schematic representation of main features of expression and reporter constructs. (A) In the expression plasmid pCMV-tetR-KRAB, the TetR-KRAB sequence is placed downstream of the strong immediate-early human CMV promoter (P_{CMV}) and is followed by an SV40 polyadenylation signal (An). The TetR-KRAB fusion protein is composed of the complete 207-amino-acid sequence of TetR fused to 6 amino acids representing the SV40 large-T-antigen nuclear localization sequence (NLS) and a 121-amino-acid KRAB domain originating from the NH_2 terminus of human Kox1 containing the KRAB A and B domains (27). The KRAB domain has been removed in the control expression plasmid pCMV-tetR. (B) The reporter plasmid ptetO7-CMV-L consists of seven *tetO* sequences placed upstream of the promoter/enhancer sequence (positions −675 to +75, +1 being the transcriptional start site) of human CMV followed by sequences encoding firefly luciferase and an SV40 poly(A) site (An). The distances of the boundaries of the *tetO*-containing fragment (−994 and −685) with respect to the transcriptional start site of the CMV promoter and the sites for insertion of the promoter(s), X (*Xho*I) and B (*Bam*HI), are indicated.

or without tetracycline for 1 day (Fig. 2). In transfections using the control vector pCMV lacking TetR sequences, similar luciferase activities were obtained from cells grown with or without tetracycline. In contrast, luciferase activity was reduced about 10-fold in transfections with pCMV-tetR-KRAB (Fig. 2, CMV-tRK). In these cells, luciferase expression was restored by addition of 0.5 μg of tetracycline per ml to the medium. No repression was observed when TetR fused to the nuclear localization sequence was expressed alone (Fig. 2, pCMV-tR), demonstrating that the repression is mediated by the KRAB domain.. Similar results were obtained for the reporter construct ptetO7-TK-L carrying the HSV TK promoter (data not shown). No differences in luciferase activities were observed for reporter constructs lacking *tetO* sequences (data not shown). This observation suggested that TetR-KRAB interfered with productive initiation at the CMV or TK promoter only when bound to *tetO* sequences. We next made stable transfectants expressing TetR-KRAB under the control of the CMV promoter. The construct used (pCMV-tetR-KRAB-hyg) contained a hygromycin resistance gene under the control of the TK promoter inserted into pCMV-tetR-KRAB (see Materials and Methods). Clones resistant to hygromycin B were assayed for expression of the TetR-KRAB protein, and the intracellular localization of the TetR-KRAB hybrid protein was visualized by indirect immunofluorescence. In one of the stable cell lines examined, A12, the TetR-KRAB protein was localized predominantly to the nucleus by indirect immunofluorescence using an anti-Kox1 immune serum (Fig. 3d and h). The staining was specific since the preimmune serum showed very little staining (Fig. 3b and f). A12 cells (Fig. 3d) showed a much stronger nuclear staining than the parental HeLa cells (Fig. 3h), as a result of the expression of the TetR-KRAB protein. The staining of HeLa nuclei (Fig. 3h) suggests that nuclear proteins that are immunologically related or identical to Kox1 are expressed in HeLa cells. Kox1 mRNA has been detected in HeLa cells (35).

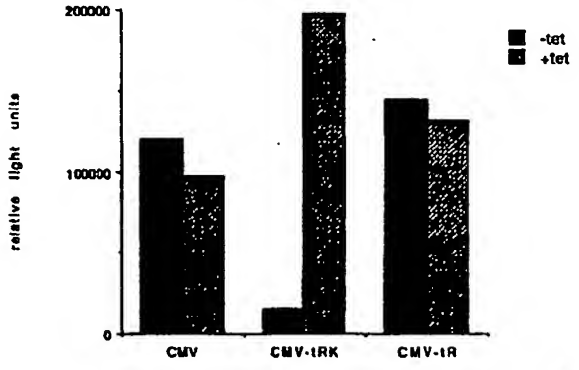


FIG. 2. Tetracycline-reversible silencing of the CMV promoter by TetR-KRAB in transiently transfected HeLa cells. HeLa cells (70 to 80% confluent) grown in 35-mm-diameter dishes without tetracycline (tet) were transiently transfected at a molar ration of 1:5 with the reporter plasmid ptetO7-CMV-L (0.5 μg) and 2.5 μg of one of the expression plasmids pCMV (CMV), pCMV-tetR-KRAB (CMV-tRK), and pCMV-tetR (CMV-tR). After transfection, cells were grown for an additional 24 h with or without 0.5 μg of tetracycline per ml before luciferase activities were determined from aliquots (1/50) of total cell extracts.

The luciferase activities from HeLa A12 cells transiently transfected with plasmids pCMV-L and pTK-L, which lack *tetO* sequences, were not altered by the presence of tetracycline in the medium (Fig. 4). However, *tetO* sequences inserted in ptetO7-CMV-L and ptetO7-TK-L resulted in a significant reduction of luciferase activities in the absence of tetracycline of more than 1 order of magnitude (Fig. 4). Addition of tetracycline fully restored the expression of luciferase (Fig. 4). Luciferase activities of these plasmids transfected into parental HeLa cells that lack stably expressed TetR-KRAB protein were not affected by tetracycline (data not shown). Increasing

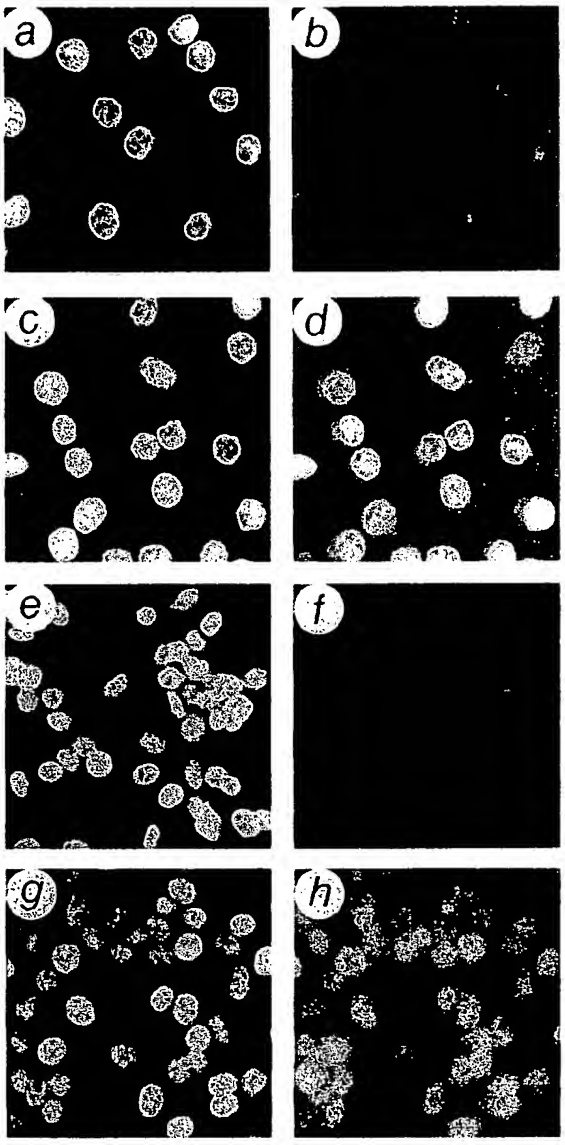


FIG. 3. Expression and nuclear localization of TetR-KRAB fusion protein in HeLa A12 cells. HeLa cells were stably transfected with pCMV-TetR-KRAB-hyg, and stable clones were selected in hygromycin B. The right-hand row of photomicrographs shows indirect immunofluorescence with an anti-Kox1 immune serum of stable transfected HeLa A12 (d) and HeLa (h) cells. A preimmune serum was used as a control for A12 (b) and HeLa (f) cells. Fluorescein-conjugated goat anti-rabbit antibodies were used as secondary antibodies. The DNAs of the respective cells were stained with DAPI to visualize the nuclear compartments and are shown in panels a, c, e, and g.

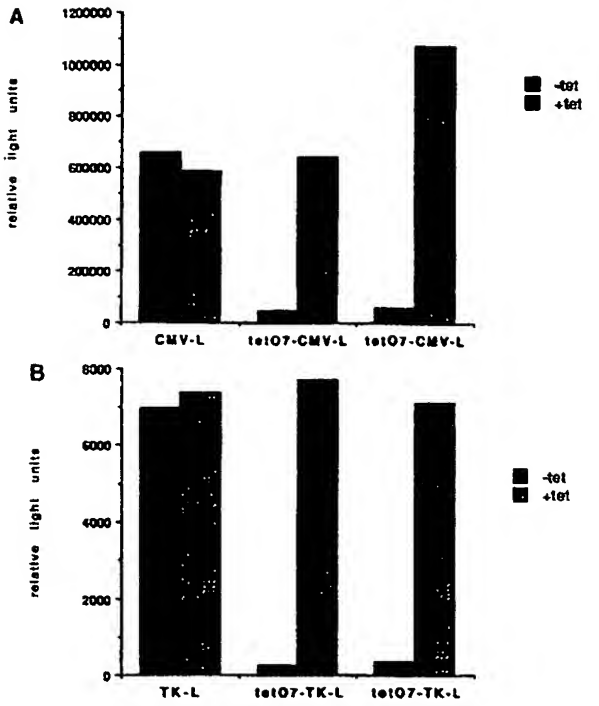


FIG. 4. Tetracycline-reversible silencing of the CMV and HSV TK promoters in HeLa A12 cells stably transfected with pCMV-tetR-KRAB. HeLa A12 cells (70 to 80% confluent) grown in 35-mm-diameter dishes without tetracycline (tet) were transiently transfected in duplicate with, per dish, 2.5 μg of pCMV-L or ptetO7-CMV-L (A) or 2.5 μg of pTK-L or ptetO7-TK-L (B). After transfection, the cells were grown in the presence or absence of 0.5 μg of tetracycline per ml in the medium for an additional 24 h before luciferase activities were determined from 1/50 of total extracts of transfected cells. Two independent experiments are shown for ptetO7-CMV-L (A) and ptetO7-TK-L (B).

the distance of the *tetO* sequences to the transcriptional start site to about 3,600 bp still resulted in significant tetracycline-reversible repression of luciferase expression by TetR-KRAB which was just threefold lower than that at a distance of 685 bp (13a). Furthermore, the orientation of the *tetO*-containing fragment with regard to the promoter had no influence on the observed repression (13a). All of the many different promoters tested so far in this system were found to respond to TetR-KRAB-mediated silencing equally well (data not shown). We conclude that TetR-KRAB stably produced in HeLa cells is able to actively repress the productive expression from both the CMV and TK promoters by binding to *cis*-acting *tetO* sequences placed at a distance from the respective transcriptional initiation sites.

Reversible repression of the ptetO7-CMV-L construct stably integrated into the genome of HeLa A12 cells. To quantify the tetracycline-inducible repression of the CMV promoter, the construct ptetO7-CMV-L was stably integrated into HeLa A12 cells that express the TetR-KRAB fusion protein. HeLa A12 cells were cotransfected with ptetO7-CMV-L linearized at the unique *Asp* 700 site and pNeo 5′ (26). One clone resistant to G418 that showed significant luciferase levels when grown in the presence of 0.5 μg of tetracycline per ml was selected and named TIS-10. Luciferase activities were determined in extracts from TIS-10 cells after growth in medium containing or lacking tetracycline. The presence of tetracycline for 24 h in the growth medium of TIS-10 cells resulted in more than 10-fold increases in luciferase activity (Fig. 5). No obvious change in growth behavior was detected in the presence of tetracycline, which demonstrates that the TetR-KRAB fusion protein can locate the *tetO* sequences in the chromatin of mammalian cells and reduce the activity of the ptetO7-CMV-L construct within the genomic environment in a reversible fashion. A tetracycline concentration of 0.1 μg/ml was sufficient to release repression (data not shown). This concentration is well below the concentration of 10 μg/ml above which inhibition of growth had been observed (reference 16 and data not shown).

Kinetics of tetracycline action. The time course of tetracy-

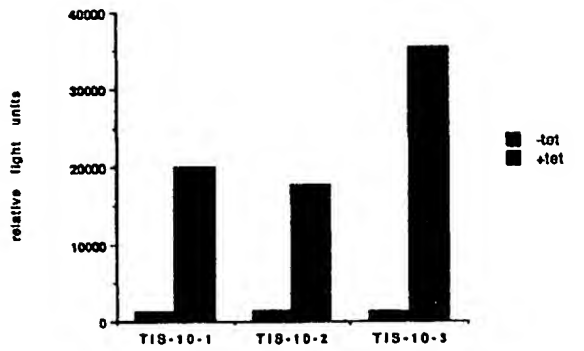


FIG. 5. Tetracycline-inducible silencing of a chromosomally integrated CMV promoter fused to *tetO* sequences. The ptetO7-CMV-L construct was stably integrated into HeLa A12 cells, resulting in the TIS-10 cell line. TIS-10 cells were grown in 35-mm-diameter dishes (70% confluence), and tetracycline (tet; 0.5 μg/ml) was either omitted or added to the growth medium. After 24 h, luciferase activities were determined from 1/50 of extracts from cells grown with or without tetracycline. Data from three individual dishes (TIS-10-1 to TIS-10-3) are shown.

cline action on luciferase expression in TIS-10 cells was analyzed in two different ways. First, tetracycline was added to cultures grown in the absence of tetracycline (Fig. 6). By 24 h, the luciferase activity was stimulated 10-fold over the noninduced values. This stimulation had increased to more than 30-fold by 2 days and more than 50-fold after 3 days. In a second experiment, we measured the decrease of luciferase activities after removal of tetracycline. Within 3 days, these levels dropped more than 50-fold (data not shown).

The hybrid protein (TetR-KRAB) produced in HeLa A12 cells is associated with a cellular protein of around 110 kDa. To investigate the mechanism of TetR-KRAB-mediated repression, we performed electrophoretic mobility shift analysis with labeled synthetic *tetO* DNA and crude extracts from A12 cells stably producing the TetR-KRAB protein. Incubation of the labeled *tetO* sequences with crude extract from A12 cells led to the formation of a very slowly migrating specific complex C1 (Fig. 7A, lane 3). The migration of this complex cannot be explained simply by a dimer of the 38-kDa TetR-KRAB protein bound to a 21-bp *tetO* DNA fragment. In several experiments, a significant fraction of the specific complexes did not even enter the gel, suggesting that large complexes might have been formed. Such complexes were not observed with extracts

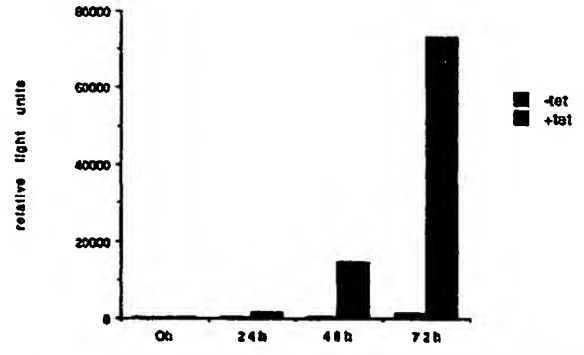


FIG. 6. Time course of tetracycline action in TIS-10 cells. TIS-10 cells were grown in the absence of tetracycline (tet), trypsinized, seeded in 35-mm-diameter dishes at low density (15% confluency), and further grown without or with 0.5 mg of tetracycline per ml in the growth medium. After the indicated times, relative luciferase activities were determined from aliquots of extracts (1/50) of cells.

from parental HeLa cells (Fig. 7A, lane 2). Complex formation was sensitive to tetracycline (1 μg/ml) and specifically competed for by an excess of unlabeled *tetO* DNA, while DNA fragments with unrelated sequences had no effect (data not shown).

To investigate further the nature of these large complexes, the ionic detergent DOC was added to the binding reactions in order to dissociate weak protein-protein interactions (2). The presence of 0.05 to 0.2% DOC resulted in the disappearance of the large complex C1 and the formation of a faster-migrating complex, C2 (Fig. 7A, lanes 4 to 6). The nature of complex C2 was further investigated by adding a Kox1-specific immune serum (28a). In the presence of a 1:200 dilution of the polyclonal anti-Kox1 immune serum together with 0.1% DOC, a retarded complex, C3, was formed (supershift; Fig. 7B, lane 5). No supershift was observed with the preimmune serum (Fig. 7B, lane 4), which suggested that the faster-migrating complex C2 was probably composed of a TetR-KRAB dimer bound to the *tetO* sequence. The large complex C1 was also detected in extracts from HeLa cells transiently transfected with pCMV-tetR-KRAB (C1; Fig. 7C, lane 3). In contrast, a much smaller complex, C4, was formed with extracts from HeLa cells transfected with pCMV-tetR expressing the 25-kDa TetR protein fused to the nuclear localization sequence but lacking the KRAB domain (C4: Fig. 7C, lane 4; see also Fig. 1 and 2). Complex C4, migrating just slightly above a nonspecific complex (asterisk in Fig. 7C), is not formed in the presence of 1 μg of tetracycline per ml and is insensitive to low levels of DOC (not shown).

The very slowly migrating complex C1 could be the result either of TetR-KRAB dimers associating with each other to form a larger complex or of a complex in which the KRAB domain of Kox1 is associated with a cellular factor(s). As a first step toward distinguishing between these possibilities, we metabolically labeled HeLa A12 cells and HeLa cells with L-[^{35}S]methionine and prepared extracts. Extracts were immunoprecipitated with the rabbit anti-Kox1 immune serum and protein G-Sepharose beads and separated by SDS-PAGE (Fig. 8). Immunoprecipitates of extracts from HeLa A12 cells showed two bands corresponding to the predicted size of the TetR-KRAB fusion protein of 38 kDa (Fig. 8, lane 2). This doublet was specifically missing in extracts from HeLa cells (Fig. 8, lane 5). The occurrence of a doublet may indicate posttranslational modification, such as phosphorylation, of the TetR-KRAB fusion protein synthesized in HeLa cells. In addition, a band corresponding to a protein of around 110 kDa was specifically coimmunoprecipitated together with TetR-KRAB (Fig. 8, SMP1, lane 2). This band was not visible in extracts from HeLa cells (Fig. 8, lane 5). Most significantly, if the protein G-Sepharose beads were extracted once with 50 μl of IP buffer containing 0.1% DOC, the 110-kDa protein was released from the beads (Fig. 8, lane 3) and appeared in the DOC supernatant (Fig. 8, lane 4). The DOC sensitivity of the association of the 110-kDa protein to TetR-KRAB in the coimmunoprecipitation is strikingly reminiscent of the DOC sensitivity of the large C1 complexes formed in extracts from HeLa A12 cells with labeled *tetO* sequences in the electrophoretic mobility shift analysis (Fig. 7A, lanes 3 to 5). In addition, the large complexes are not formed with the TetR protein lacking the KRAB domain. We therefore propose that a protein of around 110-kDa interacts with the KRAB domain of the TetR-KRAB fusion protein. This protein might be involved in the mechanism of repression of promoters from a distance by TetR-KRAB. We have tentatively named this protein SMP1.

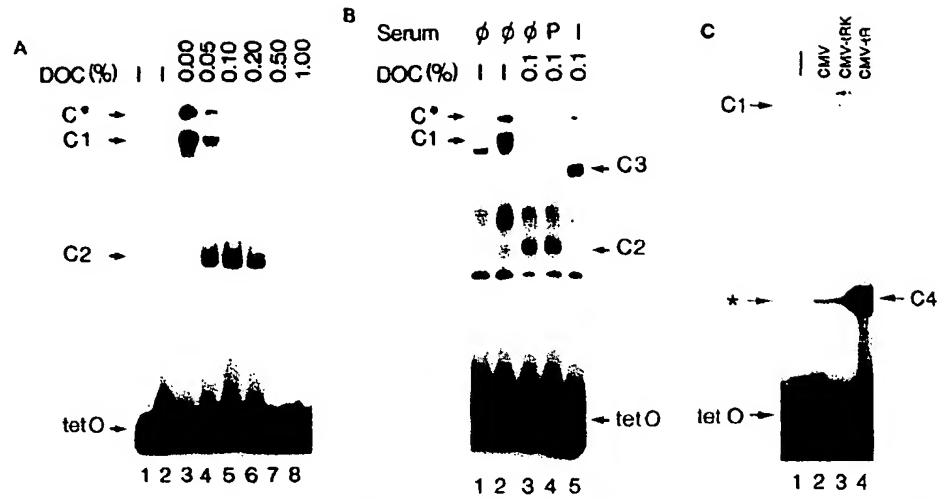


FIG. 7. DNA binding activity of TetR-KRAB from HeLa A12 cells. (A) Crude extracts from HeLa cells (lane 2) or HeLa A12 cells (lanes 3 to 8) were prepared and incubated with ^{32}P-end-labeled *tetO* DNA for 10 min before DOC was added to the indicated final concentration. After further incubation for 5 min, the samples were subjected to electrophoresis in a 4% polyacrylamide gel. As a control, no extract was added (lane 1). The positions of the free *tetO* DNA and complexes C*, C1, and C2 are indicated. (B) Crude extracts from HeLa (lane 1) or HeLa A12 (lanes 2 to 5) cells were prepared and incubated with ^{32}P-end-labeled *tetO* DNA for 10 min at room temperature before DOC was added to 0.1% (final concentration) where indicated (lanes 3 to 5). No serum (lane 3) or either a preimmune serum (P) or an anti-Kox1 immune serum (I) was added afterwards to 1% (final concentration) (lanes 4 and 5), and the samples were subjected to electrophoresis in a 4% polyacrylamide gel. The positions of the free *tetO* DNA and the different complexes (C*, C1, C2, and C3) that were formed are indicated. (C) Crude extracts from HeLa cells, transiently transfected with pCMV (lane 2), pCMV-tetR-KRAB (lane 3), or pCMV-tetR (lane 4), were prepared, and aliquots were incubated with ^{32}P-end-labeled *tetO* DNA for 15 min at room temperature and subjected to electrophoresis in a 5% polyacrylamide gel. The positions of the free *tetO* DNA, the specific complexes C1 and C4, and a nonspecific complex (*) are indicated.

DISCUSSION

In recent years, domains of various proteins that, when fused to DNA binding domains, can actively repress the activity of promoters containing appropriate target sequences have been described (reviewed in references 10 and 11). Among these is the KRAB domain of Kox1, which functions very efficiently (13a, 27). The KRAB domain is an evolutionarily conserved domain of 75 amino acids comprising a heptad repeat of methionine and leucine residues in Kox1 (35). It has been estimated that approximately one-third of all zinc finger-type DNA-binding proteins contain a KRAB domain NH_2 terminal to the zinc finger region (3). We show here that a fusion of the KRAB repression domain derived from the human Kox1 protein to the Tn*10*-derived *E. coli* TetR generated the TetR-KRAB hybrid protein, which functions as a potent DNA-binding-site-dependent transrepressor of transcription. The TetR-KRAB protein was overproduced in HeLa cells to significant levels and seems to be localized to the nucleus. The nuclear localization is presumably favored because of the presence of a nuclear localization signal derived from the SV40 large T antigen (20) between TetR and KRAB domains (Fig. 1). The strong nuclear staining in HeLa A12 cells (Fig. 3) is accompanied by the presence of a strong DNA binding activity toward *tetO* sequences in crude extracts (Fig. 7A). Furthermore, the anti-Kox1 immune serum detects the native KRAB domain in the TetR-KRAB protein, as demonstrated by the supershift of the TetR-KRAB/*tetO* complex (C3; Fig. 7B, lane 5). The TetR-KRAB produced in HeLa cells binds to *tetO* sequences in vitro, and this association is prevented by tetracycline (not shown). When the TetR-KRAB protein is bound to *tetO* sequences upstream of the enhancer and the transcriptional initiation site of promoters like the immediate-early human CMV promoter or the HSV TK promoter, the productive transcriptional activity (determined by using the luciferase gene as a reporter gene) of the respective constructs is repressed in transient assays. This repression is released by tetracycline and strictly dependent on both the presence of the KRAB domain (Fig. 2) and *cis*-acting *tetO* sequences (Fig. 4a and b). Repression by the TetR-KRAB protein is observable when *tetO* sequences are placed more than 3 kb upstream or downstream of the transcriptional initiation site of a eukaryotic promoter and independent of the orientation of the *tetO* sequences, suggest-

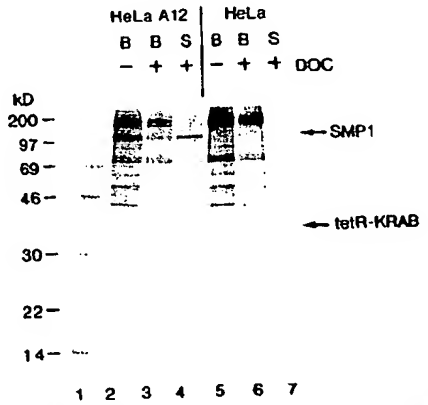


FIG. 8. A protein of around 110 kDa is coimmunoprecipitated with the TetR-KRAB from HeLa A12 cells. Whole cellular extracts of HeLa (lanes 5 to 7) and HeLa A12 (lanes 2 to 4) cells, labeled with L-[^{35}S]methionine, were incubated with an anti-Kox1 immune serum, and the immune complexes were collected with protein G-agarose beads (B; lanes 2 and 5). Half of the beads were incubated with a buffer containing 0.1% DOC, and the supernatants (S, DOC; lanes 4 and 7) were separated from the beads (B, DOC; lanes 3 and 6). The proteins on the beads and in the supernatants were analyzed by SDS-PAGE (12% gel) and fluorography. The positions of the 38-kDa TetR-KRAB fusion protein and a 110- to 120-kDa protein named SMP1 are indicated. The sizes of ^{14}C-labeled protein standards (lane 1) are indicated on the left.

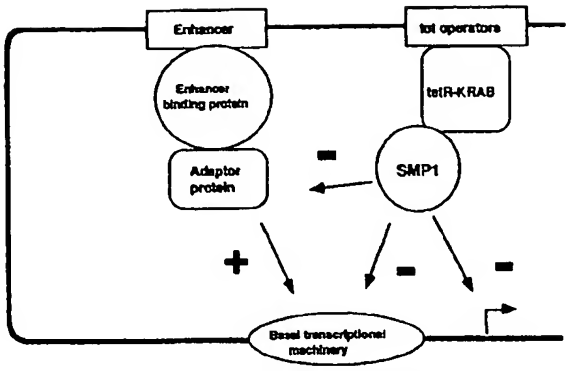


FIG. 9. Model of the silencing mechanism by TetR-KRAB and SMP1. An enhancer-binding protein exerts its stimulatory effect (+) on the basal transcriptional machinery through an adaptor protein. The TetR-KRAB protein binds to the *tetO* sequences, which can be regarded as a *cis*-acting silencer. The silencing effect (−) of the TetR-KRAB fusion protein bound to its *tetO* sequences is mediated by the SMP1 protein and could act on different levels of transcriptional control as indicated.

ing that the TetR-KRAB protein functions as a tetracycline reversible silencing protein (13a). Tetracycline-reversible silencing is also observed with TIS-10 cells that carry the ptetO7-CMV-L construct stably integrated into the genome of HeLa A12 cells expressing the TetR-KRAB protein (Fig. 5). Here the addition of tetracycline to the medium of TIS-10 cells results in a more than 50-fold activation of productive transcription from the CMV promoter (Fig. 6). Removal of tetracycline from the medium of cells grown in the presence of tetracycline restores repression (not shown). The kinetics of this process is rather slow, with half-maximal effects by 2 days. Induction or repression is observed in rapidly dividing and nondividing, i.e., confluent, cells, suggesting that replication of the DNA may not be essential for the underlying mechanism (13a).

Many other proteins contain repression domains that function when fused to heterologous DNA binding domains (8, 9, 17, 21, 24, 25, 27, 34, 37, 39). The mechanisms by which such alanine-rich or proline-rich domains, as well as the KRAB domains, exert their repressing functions have not yet been identified (10, 11, 18). We show here that the TetR-KRAB protein produced in HeLa cells forms very large complexes with *tetO* sequences in vitro (C1; Fig. 7A, lane 3). Low concentrations of DOC disrupt the large complexes and result in the formation of smaller complexes, presumably composed of TetR-KRAB dimers (C2; Fig. 7A, lanes 4 to 6). The TetR lacking the KRAB domain forms even smaller complexes (C4; Fig. 7C, lane 4). This finding suggests that at least one factor may be associated with TetR-KRAB and released upon treatment with DOC. The immunoprecipitations from extracts of HeLa A12 cells obtained by using the anti-Kox1 immune serum show that a protein of around 110 kDa is associated with the TetR-KRAB protein (Fig. 8). This protein dissociates from TetR-KRAB in the presence of DOC in a manner strikingly reminiscent of the DOC sensitivity of the large TetO/TetR-KRAB complex C1 formed with extracts from HeLa A12 cells (Fig. 7A and 8). The data suggest that the 110-kDa protein is part of the large complex C1 (Fig. 7A, lane 3). We tentatively named this protein SMP1 and propose that it is involved in mediating the silencing function of TetR-KRAB as outlined in Fig. 9. Transcriptional activator proteins bind to sites in the CMV enhancer and activate transcription from the basal transcriptional machinery in part by interacting with adaptor molecules that mediate the transactivation function; the function of SMP1 could be to mediate the silencing effect of the TetR-KRAB protein (Fig. 9). Further work will have to show whether the silencing exerted by the TetR-KRAB protein impinges directly or indirectly (i.e., via SMP1) on factors of the basal transcriptional machinery or if it works by influencing the chromatin structure of the target gene (Fig. 9). It is tempting to speculate that the KRAB-SMP1 interaction may share some similarity to protein-protein interactions that are involved in establishing repression or silencing in *Saccharomyces cerevisiae* as is the case for Ssn6-Tup1 and SIR1 (9, 21, 37).

Apart from the opportunity to understand the phenomenon of transcriptional repression exerted by the KRAB domain of Kox1, our data provide a basis for a novel way to regulate gene expression in higher mammalian cells. In contrast to a previously published system for highly efficient regulation of gene expression using a tTA-responsive promoter (15, 16), the tetracycline-reversible silencing of complex promoters by the TetR-KRAB protein offers the unique possibility of reversibly down-regulating the expression of cellular genes on top of their normal cellular regulation. Here, *tetO* sequences would be placed at a distance (0.1 to 3 kb) upstream or downstream from the transcriptional start site of the target gene such that the normal regulation is not disturbed. Binding of TetR-KRAB expressed in the very same cells results in transcriptional repression that can be released by adding tetracycline to the growth medium. Understanding the mechanism of the silencing exerted by the KRAB domain might allow improvement of the stringency of the system in order to render it applicable for the generation of cell lines or whole animals in which the expression of particular genes can be down-regulated in a reversible fashion.

ACKNOWLEDGMENTS

We thank Manfred Gossen and Hermann Bujard for plasmids pUHD15-1 and pUHC13-4 and S. K. Nordeen for pT109luc; we also thank Alexandra Livingstone and Klaus Karjalainen for helpful comments on the manuscript, Beatrice Pfeiffer, Hans Spalinger, and Hanspeter Stahlberger for artwork, and Corinne Demolliere for technical assistance.

The Basel Institute for Immunology was founded and is supported by F. Hoffmann-La Roche and Company, Ltd., Basel.

REFERENCES

1. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1987. Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
2. **Baeuerle, P. A., and D. Baltimore.** 1988. Activation of DNA-binding activity in an apparently cytoplasmic precursor of the NF-κB transcription factor. Cell 53:211–217.
3. **Bellefroid, E. J., D. A. Poncelet, P. J. Lecoq, O. Revelant, and J. A. Martial.** 1991. The evolutionarily conserved Krüppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. Proc. Natl. Acad. Sci. USA 88:3608–3612.
4. **Boshart, M., F. Weber, G. Jahn, K. Dorsch-Häsler, B. Fleckenstein, and W. Schaffner.** 1985. A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus. Cell 41:521–530.
5. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72:248–254.
6. **Braselmann, S., P. Graninger, and M. Busslinger.** 1993. A selective transcriptional induction system for mammalian cells based on Gal4-estrogen receptor fusion proteins. Proc. Natl. Acad. Sci. USA 90:1657–1661.
7. **Brent, R., and M. Ptashne.** 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. Cell 40:729–736.
8. **Bunker, C. A., and R. E. Kingston.** 1994. Transcriptional repression by *Drosophila* and mammalian polycomb group proteins in transfected mammalian cells. Mol. Cell. Biol. 14:1721–1732.
9. **Chien, C-T., S. Buck, R. Sternglanz, and D. Shore.** 1993. Targeting of SIR1 protein establishes transcriptional silencing at HM loci and telomeres in yeast. Cell 75:531–541.
10. **Clark, A. R., and K. Docherty.** 1993. Negative regulation of transcription in

eukaryotes. Biochem. J. 296:521–541.
11. **Cowell, I. G.** 1994. Repression versus activation in the control of gene transcription. Trends Biochem. Sci. 19:38–42.
12. **Deuschle, U., R. Pepperkok, F. Wang, T. J. Giordano, W. T. McAllister, W. Ansorge, and H. Bujard.** 1989. Regulated expression of foreign genes in mammalian cells under the control of coliphage T3 RNA polymerase and lac repressor. Proc. Natl. Acad. Sci. USA 86:5400–5404.
13. **Deiss, L. P., and A. Kimchi.** 1991. A genetic tool used to identify thioredoxin as a mediator of a growth inhibitory signal. Science 252:117–120.
13a. **Deuschle, U., and H.-J. Thiesen.** Unpublished data.
14. **DeWet, J. R., K. V. Wood, M. DeLuca, D. Helinski, and S. Subramani.** 1987. Firefly luciferase gene: structure and expression in mammalian cells. Mol. Cell. Biol. 7:725–737.
15. **Gossen, M., A. Bonin, and H. Bujard.** 1993. Control of gene activity in higher eukaryotic cells by prokaryotic regulatory elements. Trends Biochem. Sci. 18:471–475.
16. **Gossen, M., and H. Bujard.** 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc. Natl. Acad. Sci. USA 89:5547–5551.
17. **Han, K., and J. L. Manley.** 1993. Functional domains of the *Drosophila* engrailed protein. EMBO J. 12:2723–2733.
18. **Herschbach, B. M., and A. D. Johnson.** 1993. Transcriptional repression in eukaryotes. Annu. Rev. Cell Biol. 9:479–509.
19. **Heynes, N. E., N. Kennedy, U. Rahmsdorf, and B. Groner.** 1981. Hormone-responsive expression of an endogenous proviral gene of mammary tumour virus after molecular cloning and gene transfer into cultured cells. Proc. Natl. Acad. Sci. USA 78:2038–2042.
20. **Kalderon, D., B. L. Roberts, W. D. Richardson, and A. E. Smith.** 1984. A short amino acid sequence able to specify nuclear location. Cell 39:499–509.
21. **Keleher, C. A., M. J. Redd, J. Schultz, M. Carlson, and A. D. Johnson.** 1992. Ssn6-Tup1 is a general repressor of transcription in yeast. Cell 68:709–719.
22. **Labow, M. A., S. B. Baim, T. Shenk, and A. J. Levine.** 1990. Conversion of the *lac* repressor into an allosterically regulated transcriptional activator for mammalian cells. Mol. Cell. Biol. 10:3342–3356.
23. **Lee, F., R. Mulligan, P. Berg, and G. Ringold.** 1991. Glucocorticoids regulate expression of dihydrofolate reductase cDNA in mouse mammary tumour virus chimaeric plasmids. Nature (London) 294:228–232.
24. **Licht, J. D., W. Hanna-Rose, J. C. Reddy, M. A. English, M. Ro, M. Grossel, R. Shaknovich, and U. Hansen.** 1994. Mapping and mutagenesis of the amino-terminal transcriptional repression domain of the *Drosophila* Krüppel protein. Mol. Cell. Biol. 14:4057–4066.
25. **Licht, J. D., M. Ro, M. A. English, M. Grossel, and U. Hansen.** 1993. Selective repression of transcriptional activators at a distance by the Drosophila Krüppel protein. Proc. Natl. Acad. Sci. USA 90:11361–11365.
26. **Lusky, M., and M. R. Botchan.** 1984. Characterization of the bovine papilloma virus plasmid maintenance sequence. Cell 36:391–401.
27. **Margolin, J., J. R. Friedman, W. K.-H. Meyer, H. Vissing, H.-J. Thiesen, and F. J. Rauscher III.** 1994. Krüppel-associated boxes are potent transcriptional repression domains. Proc. Natl. Acad. Sci. USA 91:4509–4513.
28. **Mayo, K. E., R. Warren, and R. D. Palmiter.** 1982. The mouse metallothionein-I gene is transcriptionally regulated by cadmium following transfection into human or mouse cells. Cell 29:99–108.
28a. **Meyer, W. K.-H., and H.-J. Thiesen.** Unpublished data.
29. **Nordeen, S. K.** 1988. Luciferase reporter gene vectors for analysis of promoters and enhancers. BioTechniques 6:454–457.
30. **Picard, D., S. J. Salser, and K. R. Yamamoto.** 1988. A movable and regulatable inactivation function within the steroid binding domain of the glucocorticoid receptor. Cell 54:1073–1080.
31. **Rosenberg, A. H., B. N. Lade, D. S. Chui, S. W. Lin, J. L. Dunn, and F. W. Studier.** 1987. Vectors for selective expression of cloned DNA's by T7 RNA polymerase. Gene 56:125–135.
32. **Ryals, J., P. Dierks, H. Ragg, and C. Weissmann.** 1985. A 46 nucleotide promoter segment from an IFN-α gene renders an unrelated promoter inducible by virus. Cell 41:497–507.
33. **Sadowski, I., J. Ma, S. Triezenberg, and M. Ptashne.** 1988. GAL4-VP16 is an unusually potent transcriptional activator. Nature (London) 335:563–564.
34. **TenHarmsel, A., R. Austin, N. Saenelli, and M. D. Biggin.** 1993. Cooperative binding at a distance by even-skipped protein correlates with repression and suggests a mechanism of silencing. Mol. Cell. Biol. 13:2742–2752.
35. **Thiesen, H.-J.** 1990. Multiple genes encoding zinc finger domains are expressed in human T cells. New Biol. 2:363–374.
36. **Thiesen, H.-J.** 1990. Target detection assay: a general method to determine DNA binding sites for putative DNA-binding proteins. Immunol. Methods 4:61–74.
37. **Tzamarias, D., and K. Struhl.** 1994. Functional dissection of the yeast Cyc8-Tup1 transcriptional co-repressor complex. Nature (London) 369:758–761.
38. **Webster, N. J. G., S. Green, J. R. Jin, and P. Chambon.** 1988. The hormone-binding domains of the estrogen and glucocorticoid receptors contain an inducible transcription activation domain. Cell 54:199–207.
39. **Witzgall, R., E. O'Leary, A. Leaf, D. Önaldi, and J. V. Bonventre.** 1994. The Krüppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. Proc. Natl. Acad. Sci. USA 91:4514–4518.
40. **Wurm, F. M., K. A. Gwinn, and R. E. Kingston.** 1986. Inducible overexpression of the mouse c-myc protein in mammalian cells. Proc. Natl. Acad. Sci. USA 83:5414–5418.

EXHIBIT 6

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 229 134 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
07.08.2002 Bulletin 2002/32

(51) Int Cl.[7]: **C12Q 1/68**, C12N 15/10, G01N 33/50

(21) Application number: **02250681.0**

(22) Date of filing: **31.01.2002**

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE TR
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **31.01.2001 US 265805 P**
26.10.2001 US 339260 P

(71) Applicants:
- **Nucleonics, Inc**
  Malvern, Pennsylvania 19355 (US)
- **Message Pharmaceuticals, Inc.**
  Malvern, PA 19355 (US)

(72) Inventors:
- **Giordano, Tony**
  Pheonixville, PA 19460 (US)
- **Pachuk, Catherine**
  Lansdale, PA 19446 (US)
- **Satishchandran, Chandrasekhar**
  Lansdale, PA 19446 (US)

(74) Representative: **Davies, Jonathan Mark**
Reddie & Grose
16 Theobalds Road
London WC1X 8PL (GB)

(54) Use of post-transcriptional gene silencing for identifying nucleic acid sequences that modulate the function of a cell

(57) Described herein are methods for identifying nucleic acid sequences that modulate the function of a cell, the expression of a gene in a cell, or the biological activity of a target polypeptide in a cell. The methods involve the use of double stranded RNA expression libraries, double stranded RNA molecules, and post-transcriptional gene silencing techniques.

EP 1 229 134 A2

Description

Background of the Invention

[0001] The invention relates to methods for identifying nucleic acid sequences that modulate the function of a cell, by the use of post-transcriptional gene silencing.

[0002] Double stranded RNA (dsRNA) has been shown to induce sequence-specific gene silencing in a number of different organisms. Gene silencing can occur through various mechanisms, one of which is post-transcriptional gene silencing (PTGS). In post-transcriptional gene silencing, transcription of the target locus is not affected, but the RNA half-life is decreased. The mechanisms by which PTGS occurs are not yet clear. Exogenous dsRNA has been shown to act as a potent inducer of PTGS in nematodes, trypanosomes, and insects. In addition, studies in C. *elegans* and *Drosophila* show that a few molecules of dsRNA per cell are sufficient to trigger a PTGS response. Furthermore, studies in mice have demonstrated that dsRNA can interfere with the expression of genes in mouse embryos.

[0003] There exists a need to identify molecules that selectively regulate the expression of genes in vertebrate cells without the associated toxicity of the interferon response. Such regulation should allow the down-regulation of expression from genes whose gene products are detrimental to the cells.

Summary of the Invention

[0004] In general the invention features high throughput methods of using PTGS to identify a nucleic acid sequence that modulates the function of a cell, gene expression of a target nucleic acid, or the biological activity of a target polypeptide. The method involves the use of specially constructed cDNA libraries derived from a cell, for example, a primary cell or a cell line that has an observable phenotype or biological activity, (*e.g.*, an activity mediated by a target polypeptide or altered gene expression), that are transfected into cells to inhibit gene expression. This inhibition of gene expression alters the function of a cell, gene expression of a target nucleic acid, or the biological activity of a target polypeptide, and the nucleic acid sequence responsible for the modulation can be readily identified. The method may also utilize randomized nucleic acid sequences or a given sequence for which the function is not known. Although the use of PTGS as a validation strategy is known in the art, its use in screening techniques, as described herein, is novel.

[0005] Accordingly, in a first aspect, the invention features a method for identifying a nucleic acid sequence that modulates the function of a cell. The method involves: (a) transforming a population of cells with a double stranded RNA expression library, where the library is derived from the cells, where at least two cells of the population of cells are each transformed with a different nucleic acid from the double stranded RNA expression library, and where the nucleic acid is capable of forming double stranded RNA; (b)optionally selecting for a cell in which the nucleic acid is expressed in the cell; and (c) assaying for a modulation in the function of the cell, wherein a modulation identifies a nucleic acid sequence that modulates the function of a cell.

[0006] In a desirable embodiment of the first aspect of the invention, assaying for a modulation in the function of a cell comprises measuring cell motility, apoptosis, cell growth, cell invasion, vascularization, cell cycle events, cell differentiation, cell dedifferentiation, neuronal cell regeneration, or the ability of a cell to support viral replication.

[0007] In a second aspect, the invention features a method for identifying a nucleic acid sequence that modulates expression of a target nucleic acid in a cell. The method involves: (a) transforming a population of cells with a double stranded RNA expression library, where the library is derived from the cells, where at least two cells of the population of cells are each transformed with a different nucleic acid from the double stranded RNA expression library, and where the nucleic acid is capable of forming double stranded RNA; (b)optionally selecting for a cell in which the nucleic acid is expressed in the cell; and (c) assaying for a modulation in the expression of a gene in the cell, where a modulation identifies a nucleic acid sequence that modulates expression of a target nucleic acid in a cell.

[0008] In a desirable embodiment of the second aspect of the invention, the target nucleic acid is assayed using DNA array technology.

[0009] In a third aspect, the invention features a method for identifying a nucleic acid sequence that modulates the biological activity of a target polypeptide in a cell. The method involves: (a) transforming a population of cells with a double stranded RNA expression library, where the library is derived from the cells, where at least two cells of the population of cells are each transformed with a different nucleic acid from the double stranded RNA expression library, and where the nucleic acid is capable of forming double stranded RNA; (b)optionally selecting for a cell in which the nucleic acid is expressed in the cell; and (c) assaying for a modulation in the biological activity of a target polypeptide in the cell, wherein a modulation identifies a nucleic acid sequence that modulates the biological activity of a target polypeptide.

[0010] In one embodiment of any of the above aspects of the invention, in transforming step (a), the nucleic acid is stably integrated into a chromosome of the cell. Integration of the nucleic acid may be random or site-specific. Desirably integration is mediated by recombination or retroviral insertion. In addition, desirably a single copy of the nucleic acid is integrated into the chromosome. In another embodiment of any of the above aspects of the invention, in step (a) at least 50, more desirably 100; 500; 1000; 10,000; or 50,000 cells of the population of cells are each transformed with a

different nucleic acid from the double stranded RNA expression library. In other embodiments, the population of cells is transformed with at least 5%, more desirably at least 25%, 50%, 75%, or 90%, and most desirably at least 95% of the double stranded RNA expression library. In yet another embodiment, the method further involves: (d) identifying the nucleic acid sequence by amplifying and cloning the sequence. Desirably amplification of the sequence involves the use of the polymerase chain reaction (PCR).

[0011] In other embodiments of any of the above aspects of the invention, the double stranded RNA expression library contains cDNAs or randomized nucleic acids. The double stranded RNA expression library may be a nuclear double stranded RNA expression library, in which case the double stranded nucleic acid is made in the nucleus. Alternatively, the double stranded RNA expression library may be a cytoplasmic double stranded RNA expression library, in which case the double stranded nucleic acid is made in the cytoplasm. In addition, the nucleic acid from the double stranded RNA expression library may be made *in vitro* or *in vivo*. In addition, the identified nucleic acid sequence may be located in the cytoplasm of the cell.

In still another embodiment of any of the above aspects of the invention, the nucleic acid is contained in a vector, for example a double stranded RNA expression vector. The vector may then be transformed such that it is stably integrated into a chromosome of the cell, or it may function as an episomal (non-integrated) expression vector within the cell. In one embodiment, a vector that is integrated into a chromosome of the cell contains a promoter operably linked to a nucleic acid encoding ahairpin or double stranded RNA. In another embodiment, the vector does not contain a promoter operably linked to a nucleic acid encoding a double stranded RNA. In thislater embodiment, the vector integrates into a chromosome of a cell such that an endogenous promoter is operably linked to a nucleic acid from the vector that encodes a double stranded RNA. Desirably, the double stranded RNA expression vector comprises at least one RNA polymerase II promoter, for example, a human CMV-immediate early promoter (HCMV-IE) or a simian CMV (SCMV) promoter, at least one RNA polymerase I promoter, or at least one RNA polymerase III promoter. The promoter may also be a T7 promoter, in which case, the cell further comprises T7 polymerase. Alternatively, the promoter may be an SP6 promoter, in which case, the cell further comprises SP6 polymerase. The promoter may also be one convergent T7 promoter and one convergent SP6 promoter. A cell may be made to contain T7 or SP6 polymerase by transforming the cell with a T7 polymerase or an SP6 polymerase expression plasmid, respectively. In some embodiments, a T7 promoter or a RNA polymerase III promoter is operably linked to a nucleic acid that encodes a small double stranded RNA (*e.g.*, a double stranded RNA that is less than 200, 150, 100, 75, 50, or 25 nucleotides in length). In other embodiments, the promoter is a mitochondrial promoter that allows cytoplasmic transcription of the nucleic acid in the vector (see, for example, the mitochondrial promoters described in WO 00/63364, filed April 19, 2000). Alternatively, the promoter is an inducible promoter, such as a *lac (Cronin et al. Genes & Development 15: 1506-1517, 2001), ara (Khlebnikov et al.,* J Bacteriol. 2000 Dec;182(24):7029-34), ecdysone (Rheogene, www.rheogene.com), RU48 (mefepristone) (corticosteroid antagonist) (Wang XJ, Liefer KM, Tsai S, O'Malley BW, Roop DR, Proc Natl Acad Sci U S A. 1999 Jul 20; 96(15):8483-8), or *tet* promoter (Rendal *et al.*, Hum Gene Ther. 2002 Jan;13(2):335-42. and Larnartina *et al.*, Hum Gene Ther. 2002 Jan;13(2):199-210) or a promoter disclosed in WO 00/63364, filed April 19, 2000. In desirable embodiments, the inducible promoter is not induced until all the episomal vectors are eliminated from the cell. The vector may also comprise a selectable marker. In addition, these vectors may be used in combination with methods that inhibit or prevent an interferon response or double stranded RNA stress response, as described herein.

[0012] Desirably in a vector for use in any of the above aspects of the invention, the sense strand and the antisense strand of the nucleic acid sequence are transcribed from the same nucleic acid sequence using two convergent promoters. In another desirable embodiment, in a vector for use in any of the above aspects of the invention, the nucleic acid sequence comprises an inverted repeat, such that upon transcription, the nucleic acid forms a double stranded RNA.

[0013] In still other embodiments of any of the above aspects of the invention, the cell and the vector each further comprise a *loxP* site and site-specific integration of the nucleic acid into a chromosome of the cell occurs through recombination between the *loxP* sites. In addition, step (b) of any of the above aspects of the invention further involves rescuing the nucleic acid through Cre-mediated double recombination.

[0014] In still further embodiments of any of the above aspects of the invention, the identified nucleic acid sequence is located in the nucleus of the cell. Alternatively, the identified nucleic acid sequence may be located in the cytoplasm of the cell.

[0015] In yet another embodiment of any of the above aspects of the invention, the nucleic acid from the double stranded RNA expression library is at least 100, 500, 600, or 1000 nucleotides in length. In other embodiments of any of the above aspects of the invention, the nucleic acid from the double stranded RNA expression library is at least 10, 20, 30, 40, 50, 60, 70, 80, or 90 nucleotides in length. In yet other embodiments, the number of nucleotides in the nucleic acid from the double stranded RNA expression library is between 5-100 nucleotides, 15-100 nucleotides, 20-95 nucleotides, 25-90 nucleotides, 35-85 nucleotides, 45-80 nucleotides, 50-75 nucleotides, or 55-70 nucleotides, inclusive. In still other embodiments, the number of nucle-

otides in the nucleic acid from the double stranded RNA expression library is contained in one of the following ranges: 5-15 nucleotides, 15-20 nucleotides, 20-25 nucleotides, 25-35 nucleotides, 35-45 nucleotides, 45-60 nucleotides, 60-70 nucleotides, 70-80 nucleotides, 80-90 nucleotides, or 90-100 nucleotides, inclusive. In other embodiments, the nucleic acid contains less than 50,000; 10,000; 5,000; or 2,000 nucleotides. In addition, the nucleic acid from the double stranded RNA expression library may contain a sequence that is less than a full length RNA sequence.

[0016] In still further embodiments of any of the above aspects of the invention, the cell is a plant cell or an animal cell. Desirably the animal cell is a vertebrate or mammalian cell, for example, a human cell. The cell may be *ex vivo* or *in vivo*. The cell may be a gamete or a somatic cell, for example, a cancer cell, a stem cell, a cell of the immune system, a neuronal cell, a muscle cell, or an adipocyte.

[0017] Transformation/transfection of the cell may occur through a variety of means including, but not limited to, lipofection, DEAE-dextran-mediated transfection, microinjection, protoplast fusion, calcium phosphate precipitation, viral or retroviral delivery, electroporation, or biolistic transformation. The RNA or RNA expression vector (DNA) may be naked RNA or DNA or local anesthetic complexed RNA or DNA (Pachuk *et al.*, Biochim. Biophys. Acta 1468:20-30, 2000). In another embodiment, the cell is not a *C. elegans* cell. Desirably the vertebrate or mammalian cell has been cultured for only a small number of passages (*e.g.*, less than 30 passages of a cell line that has been directly obtained from American Type Culture Collection), or are primary cells. Desirably, the vertebrate or mammalian cell is transformed with nucleic acids that are not complexed with cationic lipids.

[0018] In yet another embodiment of any of the above aspects of the invention, the cell is derived from a parent cell, and is generated by: (a) transforming a population of parent cells with a bicistronic plasmid expressing a selectable marker and a reporter gene, and comprising a *loxP* site; (b) selecting for a cell in which the plasmid is stably integrated; and (c) selecting for a cell in which one copy of the plasmid is stably integrated in a transcriptionally active locus. Desirably the selectable marker is G418 and the reporter gene is green fluorescent protein (GFP).

[0019] In still another embodiment of the above aspects of the invention, generation of the double stranded expression library comprises: (a) isolating RNA from a cell; (b) synthesizing cDNAs from the RNA of step (a); and (c) cloning each cDNA into a vector. Desirably cDNA synthesis is optimized and/or size selected for the generation and/or selection of cDNAs that are at least 100, 500, 600, or 1000 nucleotides in length. In other embodiments, the cDNAs are least 10, 20, 30, 40, 50, 60, 70, 80, or 90 nucleotides in length. In yet other embodiments, the number of nucleotides in the cDNAs is between 5-100 nucleotides, 15-100 nucleotides, 20-95 nucleotides, 25-90 nucleotides, 35-85 nucleotides, 45-80 nucleotides, 50-75 nucleotides, or 55-70 nucleotides, inclusive. In still other embodiments, the number of nucleotides in the cDNAs is contained in one of the following ranges: 5-15 nucleotides, 15-20 nucleotides, 20-25 nucleotides, 25-35 nucleotides, 35-45 nucleotides, 45-60 nucleotides, 60-70 nucleotides, 70-80 nucleotides, 80-90 nucleotides, or 90-100 nucleotides, inclusive. In other embodiments, the cDNAs contain less than 50,000; 10,000; 5,000; or 2,000 nucleotides. In addition, the cDNA may encode an RNA fragment that is less than full length. Desirably the vector comprises two convergent T7 promoters, two convergent SP6 promoters, or one convergent T7 promoter and one convergent SP6 promoter, a selectable marker, and/or a *loxP* site.

[0020] In an additional embodiment of any of the above aspects of the invention, the method is carried out under conditions that inhibit or prevent an interferon response or double stranded RNA stress response.

[0021] In a fourth aspect, the invention features a method for identifying a nucleic acid sequence that modulates the function of a cell, involving: (a) transforming a population of cells with a double stranded RNA that is derived from the cells; (b)optionally selecting for a cell in which the nucleic acid is expressed; and (c) assaying for a modulation in the function of the cell, wherein the modulation identifies a nucleic acid sequence that modulates the function of a cell, wherein the method is desirably carried out under conditions that inhibit or prevent an interferon response or double stranded RNA stress response.

[0022] In a desirable embodiment of the fourth aspect of the invention, assaying for a modulation in the function of a cell comprises measuring cell motility, apoptosis, cell growth, cell invasion, vascularization, cell cycle events, cell differentiation, cell dedifferentiation, neuronal cell regeneration, or the ability of a cell to support viral replication.

[0023] In a fifth aspect, the invention features a method for identifying a nucleic acid sequence that modulates expression of a target nucleic acid in a cell, involving: (a) transforming a population of cells with a double stranded RNA that is derived from the cells; (b)optionally selecting for a cell in which the nucleic acid is expressed; and (c) assaying for a modulation in the expression of the gene in the cell, wherein the modulation identifies a nucleic acid sequence that modulates expression of a target nucleic acid in a cell, wherein the method is desirably carried out under conditions that inhibit or prevent an interferon response or double stranded RNA stress response.

[0024] In a desirable embodiment of the fifth aspect of the invention, the target nucleic acid is assayed using DNA array technology.

[0025] In a sixth aspect, the invention features a method for identifying a nucleic acid sequence that modulates the biological activity of a target polypeptide in a

cell, involving: (a) transforming a population of cells with a double stranded RNA that is derived from the cells; (b)optionally selecting for a cell in which the nucleic acid is expressed in the cell; and (c) assaying for a modulation in the biological activity of a target polypeptide in the cell, wherein the modulation identifies a nucleic acid sequence that modulates the biological activity of a target polypeptide in a cell, wherein the method is desirably carried out under conditions that inhibit or prevent an interferon response or double stranded RNA stress response.

[0026] In a seventh aspect, the invention features a method for identifying a nucleic acid sequence that modulates the function of a cell, involving: (a) transforming a population of cells with a double stranded RNA; (b) optionally selecting for a cell in which the nucleic acid is expressed; and (c) assaying for a modulation in the function of the cell. Desirably, the modulation identifies a nucleic acid sequence that modulates the function of a cell, wherein the method is desirably carried out under conditions that or prevent an interferon response or double stranded RNA stress response.

[0027] In a desirable embodiment of the seventh aspect of the invention, assaying for a modulation in the function of a cell comprises measuring cell motility, apoptosis, cell growth, cell invasion, vascularization, cell cycle events, cell differentiation, cell dedifferentiation, neuronal cell regeneration, or the ability of a cell to support viral replication.

[0028] In a eighth aspect, the invention features a method for identifying a nucleic acid sequence that modulates expression of a target nucleic acid in a cell, involving: (a) transforming a population of cells with a double stranded RNA; (b)optionally selecting for a cell in which the nucleic acid is expressed; and (c) assaying for a modulation in the expression of the gene in the cell, wherein the modulation identifies a nucleic acid sequence that modulates expression of a target nucleic acid in a cell. Desirably, the method is carried out under conditions that inhibit or prevent an interferon response or double stranded RNA stress response.

[0029] In a desirable embodiment of the eighth aspect of the invention, the target nucleic acid is assayed using DNA array technology.

[0030] In a ninth aspect, the invention features a method for identifying a nucleic acid sequence that modulates the biological activity of a target polypeptide in a cell, involving: (a) transforming a population of cells with a double stranded RNA; (b)optionally selecting for a cell in which the nucleic acid is expressed in the cell; and (c) assaying for a modulation in the biological activity of a target polypeptide in the cell, wherein the modulation identifies a nucleic acid sequence that modulates the biological activity of a target polypeptide in a cell. Desirably, the method is carried out under conditions that inhibit or prevent an interferon response double stranded RNA stress response.

[0031] In one embodiment of any of the above aspects of the invention, in step (a) at least 2, more desirably 50; 100; 500; 1000; 10,000; or 50,000 cells of the population of cells are each transformed with a different double stranded RNA from a double stranded RNA expression library. Desirably, at most one double stranded RNA is inserted into each cell. In other embodiments, the population of cells is transformed with at least 5%, more desirably at least 25%, 50%, 75%, or 90%, and most desirably, at least 95% of the double stranded RNA expression library. In still another embodiment of any of the fourth, fifth, or sixth aspects of the invention, the method further involves: (d) identifying the nucleic acid sequence by amplifying and cloning the sequence. Desirably amplification of the sequence involves the use of the polymerase chain reaction (PCR).

[0032] In a tenth aspect, the invention features a cell or a population of cells that expresses a double stranded RNA that (i) modulates a function of the cell, (ii) modulates the expression of a target nucleic acid (*e.g.*, an endogenous or pathogen gene) in the cell, and/or (iii) modulates the biological activity of a target protein (*e.g.*, an endogenous or pathogen protein) in the cell. Desirably, the cell contains only one molecular species of double stranded RNA or only one copy of a double stranded RNA expression vector (*e.g.*, a stably integrated vector). Desirably, the cell or population of cells is produced using one or more methods of the invention. In other embodiments, the double stranded RNA is expressed under conditions that inhibit or prevent an interferon response or a double stranded RNA stress response

[0033] In other embodiments of any of the fourth, fifth, sixth, seventh, eighth, ninth, or tenth aspects of the invention, the double stranded RNA is derived from cDNAs or randomized nucleic acids. In addition, the double stranded RNA may be a cytoplasmic double stranded RNA, in which case the double stranded nucleic acid is made in the cytoplasm. The double stranded RNA may be made *in vitro* or *in vivo*. In addition, the identified nucleic acid sequence may be located in the cytoplasm of the cell.

[0034] In still another embodiment of any of the fourth, fifth, sixth, seventh, eighth, ninth, or tenth aspects of the invention, the nucleic acid is contained in a vector, for example, a double stranded RNA expression vector that is capable of forming a double stranded RNA. Desirably the double stranded RNA expression vector comprises at least one promoter. The promoter may be a T7 promoter, in which case, the cell further comprises T7 polymerase. Alternatively, the promoter may be an SP6 promoter, in which case, the cell further comprises SP6 polymerase. The promoter may also be one convergent T7 promoter and one convergent SP6 promoter. A cell may be made to contain T7 or SP6 polymerase by transforming the cell with a T7 polymerase or an SP6 polymerase expression plasmid, respectively. The vector may also comprise a selectable marker, for example hygromycin.

[0035] Desirably in a vector for use in any of the fourth,

fifth, sixth, seventh, eighth, ninth, or tenth aspects of the invention, the sense strand and the antisense strand of the nucleic acid sequence are transcribed from the same nucleic acid sequence using two convergent promoters. In another desirable embodiment, in a vector for use in any of the above aspects of the invention, the nucleic acid sequence comprises an inverted repeat, such that upon transcription, the nucleic acid forms a double stranded RNA.

[0036] In yet another embodiment of any of the fourth, fifth, sixth, seventh, eighth, ninth, or tenth aspects of the invention, the double stranded RNA is at least 100, 500, 600, or 1000 nucleotides in length. In other embodiments of any of the fourth, fifth, or sixth aspects of the invention, the double stranded RNA is at least 10, 20, 30, 40, 50, 60, 70, 80, or 90 nucleotides in length. In yet other embodiments, the number of nucleotides in the double stranded RNA is between 5-100 nucleotides, 15-100 nucleotides, 20-95 nucleotides, 25-90 nucleotides, 35-85 nucleotides, 45-80 nucleotides, 50-75 nucleotides, or 55-70 nucleotides, inclusive. In still other embodiments, the number of nucleotides in the double stranded RNA is contained in one of the following ranges: 5-15 nucleotides, 15-20 nucleotides, 20-25 nucleotides, 25-35 nucleotides, 35-45 nucleotides, 45-60 nucleotides, 60-70 nucleotides, 70-80 nucleotides, 80-90 nucleotides, or 90-100 nucleotides, inclusive. In other embodiments, the double stranded RNA contains less than 50,000; 10,000; 5,000; or 2,000 nucleotides. In addition, the double stranded RNA may contain a sequence that is less than a full length RNA sequence.

[0037] In still further embodiments of any of the fourth, fifth, sixth, seventh, eighth, ninth, or tenth aspects of the invention, the cell is a plant cell or an animal cell. Desirably the animal cell is a vertebrate or mammalian cell, for example, a human cell. The cell may be *ex vivo* or *in vivo*. The cell may be a gamete or a somatic cell, for example, a cancer cell, a stem cell, a cell of the immune system, a neuronal cell, a muscle cell, or an adipocyte.

[0038] In other embodiments of any of the first, second, third, seventh, eighth, ninth, or tenth aspects of the invention, the double stranded RNA is derived from a cell or a population of cells and used to transform another cell population of either the same cell type or a different cell type. In desirable embodiments, the transformed cell population contains cells of a cell type that is related to the cell type of the cells from which the double stranded RNA was derived (*e.g.*, the transformation of cells of one neuronal cell type with the double stranded RNA derived from cells of another neuronal cell type). In yet other embodiments of any of these aspects, the double stranded RNA contains one or more contiguous or non-contiguous positions that are randomized (*e.g.*, by chemical or enzymatic synthesis using a mixture of nucleotides that may be added at the randomized position). In still other embodiments, the double stranded RNA is a randomized nucleic acid in which segments of ribonucleotides and/or deoxyribonucleotides are ligated to form the double stranded RNA.

[0039] In other embodiments of any of various aspects of the invention, the double stranded RNA specifically hybridizes to a target nucleic acid but does not substantially hybridize to non-target molecules, which include other nucleic acids in the cell or biological sample having a sequence that is less than 99, 95, 90, 80, or 70% identical or complementary to that of the target nucleic acid. Desirably, the amount of the these non-target molecules hybridized to, or associated with, the double stranded RNA, as measured using standard assays, is 2-fold, desirably 5-fold, more desirably 10-fold, and most desirably 50-fold lower than the amount of the target nucleic acid hybridized to, or associated with, the double stranded RNA. In other embodiments, the amount of a target nucleic acid hybridized to, or associated with, the double stranded RNA, as measured using standard assays, is 2-fold, desirably 5-fold, more desirably 10-fold, and most desirably 50-fold greater than the amount of a control nucleic acid hybridized to, or associated with, the double stranded RNA. Desirably, the double stranded RNA only hybridizes to one target nucleic acid from a cell under denaturing, high stringency hybridization conditions. In certain embodiments, the double stranded RNA is substantially homologous (*e.g.*, at least 80, 90, 95, 98,or 100% homologous) to only one target nucleic acid from a cell. In other embodiments, the double stranded RNA is homologous to multiple RNAs, such as RNAs from the same gene family. In yet other embodiments, the double stranded RNA is homologous to distinctly different mRNA sequences from genes that are similarly regulated (*e.g.*, developmental, chromatin remodeling, or stress response induced). In other embodiments, the double stranded RNA is homologous to a large number of RNA molecules, such as a double stranded RNA designed to induce a stress response or apoptosis. In other embodiments, the percent decrease in the expression of a target nucleic acid is at least 2, 5, 10, 20, or 50 fold greater than the percent decrease in the expression of a non-target or control nucleic acid. Desirably, the double stranded RNA inhibits the expression of a target nucleic acid but has negligible, if any, effect on the expression of other nucleic acids in the cell. Examples of control nucleic acids include nucleic acids with a random sequence or nucleic acids known to have little, if any, affinity for the double stranded RNA.

[0040] In other embodiments of any of various aspects of the invention, at most one molecular species of double stranded RNA is inserted into each cell. In other embodiments, at most one vector is stably integrated into the genome of each cell. In various embodiments, the double stranded RNA is active in the nucleus of the transformed cell and/or is active in the cytoplasm of the transformed cell. In various embodiments, at least 1, 10, 20, 50, 100, 500, or 1000 cells or all of the cells in the population are selected as cells that contain or express a double stranded RNA. In some embodiments, at least

1, 10, 20, 50, 100, 500, or 1000 cells or all of the cells in the population are assayed for a modulation in the function of the cell, a modulation in the expression of a target nucleic acid (*e.g.*, an endogenous or pathogen gene) in the cell, and/or a modulation in the biological activity of a target protein (*e.g.*, an endogenous or pathogen protein) in the cell.

[0041] In other embodiments, the double stranded RNA or double stranded RNA expression vector is complexed with one or more cationic lipids or cationic amphiphiles, such as the compositions disclosed in US 4,897,355 (Eppstein *etal.*, filed October 29, 1987), US 5,264,618 (Felgner *et al.*, filed April 16, 1991) or US 5,459,127 (Felgner *et al.*, filed September 16, 1993). In other embodiments, the double stranded RNA or double stranded RNA expression vector is complexed with a liposomes/liposomic composition that includes a cationic lipid and optionally includes another component such as a neutral lipid (see, for example, US 5,279,833 (Rose), US 5,283,185 (Epand), and US 5,932,241). In yet other embodiments, the double stranded RNA or double stranded RNA expression vector is complexed with any other composition that is devised by one of ordinary skill in the fields of pharmaceutics and molecular biology.

[0042] Desirably, the double stranded RNA specifically hybridizes to a target nucleic acid but does not substantially hybridize to non-target molecules, which include other nucleic acids in the cell or biological sample having a sequence that is less than 99, 95, 90, 80, or 70 % identical to or complementary to that of the target nucleic acid. In other embodiments, the percent decrease in the expression of a target nucleic acid is at least 2, 5, 10, 20, or 50 fold greater than the percent decrease in the expression of a non-target or control nucleic acid. Desirably, the double stranded RNA inhibits the expression of the target nucleic acid but has negligible, if any, effect on the expression of other nucleic acids in the cell.

[0043] Transformation/transfection of the cell may occur through a variety of means including, but not limited to, lipofection, DEAE-dextran-mediated transfection, microinjection, protoplast fusion, calcium phosphate precipitation, viral or retroviral delivery, electroporation, or biolistic transformation. The RNA or RNA expression vector (DNA) may be naked RNA or DNA or local anesthetic complexed RNA or DNA (Pachuk *et al.*, *supra*). In yet another embodiment, the cell is not a *C. elegans* cell. Desirably the vertebrate or mammalian cell has been cultured for only a small number of passages (*e.g.*, less than 30 passages of a cell line that has been directly obtained from American Type Culture Collection), or are primary cells. In addition, desirably the vertebrate or mammalian cell is transformed with double stranded RNA that is not complexed with cationic lipids.

[0044] The transcription systems described herein provide advantages to other double stranded expression systems. Following transformation of the double stranded RNA library, cells contain hundreds to thousands of double stranded RNA expression cassettes, with concomitant expression of that many expression cassettes. In the double stranded RNA expression system of the present invention, double stranded RNA (dsRNA) expression cassettes contained within the expression vector integrate into the chromosome of the transfected cell. Desirably, every transformed cell integrates one of the double stranded expression cassettes. Through expansion of the transformed cell, episomal (non-integrated) expression vectors are diluted out of the cell over time. Desirably no transcription occurs until the episomal expression vectors are diluted out of the cell, such that not more than 5 episomal vectors remain in the cell. Most desirably, no transcription occurs until all of the episomal vectors have been diluted out of the cell, and only the integrated expression cassette remains. The time it takes for all episomal vectors to be removed from the cell is proportional to the replication rate of the transformed cell, and is generally on the order of two to several weeks of cell culture and growth. The numbers of copies of a dsRNA molecule in a transformed cell can be determined using, for example, standard PCR techniques, and thereby, the number of episomal vectors in a given cell can be monitored.

[0045] Once a stable integrant containing five or fewer, and desirably no episomal expression vectors, transcription is induced, allowing dsRNA to be expressed in the cells. This method ensures that, if desired, only one species or not more than about five species of dsRNA is expressed per cell, as opposed to other methods that express hundreds to thousands of double stranded species.

[0046] Another problem that can occur in other double stranded expression systems or dsRNA delivery systems is that some dsRNA sequences, possibly in certain cell types and through certain delivery methods, may result in an interferon response (Jaramillo *et al.*, Cancer Invest. 13:327-338, 1995). During the induction of post-transcriptional gene silencing events, induction of an interferon response is not desired, as this could lead to cell death and possibly to the prevention of gene silencing. An additional advantage of the present invention is that the dsRNA delivery methods described herein are performed such that an interferon response is inhibited or prevented.

[0047] One of the components of an interferon response is the induction of the interferon-induced protein kinase PKR (Jaramillo *et al.*, *supra*). Suppression of the interferon response and/or the PKR response, using techniques described herein, is desired in the cells targeted for a PTGS event in those instances where an interferon response would otherwise be induced. Methods for suppressing an interferon response or dsRNA stress response can be used in combination with any of the methods for identifying a nucleic acid sequence that modulates the function of a cell, gene expression in a cell, or the biological activity of a target polypeptide.

[0048] The methods of the present invention provide a means for high throughput identification of nucleic acid sequences involved in modulating the function of a cell, the expression of a target nucleic acid in a cell, or the biological activity of a target polypeptide in a cell. By transforming a population of cells with a double stranded RNA expression library, the effects of many PTGS events on cell function, expression of a target nucleic acid in a cell, or the biological activity of a target polypeptide in a cell can be evaluated simultaneously, thereby allowing for rapid identification of the nucleic acid sequence involved in a cell function, target nucleic acid expression, or biological activity of a target polypeptide of interest.

[0049] By "nucleic acid," "nucleic acid sequence," "double stranded RNA nucleic acid sequence," or "double stranded RNA nucleic acid" is meant a nucleic acid or a portion thereof that is free of the genes that, in the naturally-occurring genome of the organism from which the nucleic acid sequence of the invention is derived, flank the gene. The term therefore includes, for example, a recombinant DNA, with or without 5' or 3' flanking sequences that is incorporated into a vector, for example, a double stranded RNA expression vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote; or which exists as a separate molecule (*e.g.*, a cDNA or a genomic or cDNA fragment produced by PCR or restriction endonuclease digestion) independent of other sequences.

[0050] By "double stranded RNA" is meant a nucleic acid containing a region of two or more nucleotides that are in a double stranded conformation. In various embodiments, the double stranded RNA consists entirely of ribonucleotides or consists of a mixture of ribonucleotides and deoxynucleotides, such as the RNA/DNA hybrids disclosed, for example, by WO 00/63364, filed April 19, 2000 or U.S.S.N. 60/130,377, filed April 21, 1999. The double stranded RNA may be a single molecule with a region of self-complimentarity such that nucleotides in one segment of the molecule base pair with nucleotides in another segment of the molecule. In various embodiments, a double stranded RNA that consists of a single molecule consists entirely of ribonucleotides or includes a region of ribonucleotides that is complimentary to a region of deoxyribonucleotides. Alternatively, the double stranded RNA may include two different strands that have a region of complimentarity to each other. In various embodiments, both strands consist entirely of ribonucleotides, one strand consists entirely of ribonucleotides and one strand consists entirely of deoxyribonucleotides, or one or both strands contain a mixture of ribonucleotides and deoxyribonucleotides. Desirably, the regions of complimentarity are at least 70, 80, 90, 95, 98, or 100% complimentary. Desirably, the region of the double stranded RNA that is present in a double stranded conformation includes at least 5, 10, 20, 30, 50, 75,100, 200, 500, 1000, 2000 or 5000 nucleotides or includes all of the nucleotides in a cDNA being represented in the double stranded RNA. In some embodiments, the double stranded RNA does not contain any single stranded regions. such as single stranded ends, or the double stranded RNA is a hairpin. Desirable RNA/DNA hybrids include a DNA strand or region that is an antisense strand or region (*e.g*, has at least 70, 80, 90, 95, 98, or 100% complimentary to a target nucleic acid) and an RNA strand or region that is an sense strand or region (*e.g*, has at least 70, 80, 90, 95, 98, or 100% identity to a target nucleic acid). In various embodiments, the RNA/DNA hybrid is made *in vitro* using enzymatic or chemical synthetic methods such as those described herein or those described in WO 00/63364, filed April 19, 2000 or U.S.S.N. 60/130,377, filed April 21, 1999. In other embodiments, a DNA strand synthesized *in vitro* is complexed with an RNA strand made *in vivo* or *in vitro* before, after, or concurrent with the transformation of the DNA strand into the cell. In yet other embodiments, the double stranded RNA is a single circular nucleic acid containing a sense and an antisense region, or the double stranded RNA includes a circular nucleic acid and either a second circular nucleic acid or a linear nucleic acid (see, for example, WO 00/63364, filed April 19, 2000 or U.S.S.N. 60/130,377, filed April 21, 1999.) Exemplary circular nucleic acids include lariat structures in which the free 5' phosphoryl group of a nucleotide becomes linked to the 2' hydroxyl group of another nucleotide in a loop back fashion.

[0051] In other embodiments, the double stranded RNA includes one or more modified nucleotides in which the 2' position in the sugar contains a halogen (such as flourine group) or contains an alkoxy group (such as a methoxy group) which increases the half-life of the double stranded RNA *in vitro* or *in vivo* compared to the corresponding double stranded RNA in which the corresponding 2' position contains a hydrogen or an hydroxyl group. In yet other embodiments, the double stranded RNA includes one or more linkages between adjacent nucleotides other than a naturally-occurring phosphodiester linkage. Examples of such linkages include phosphoramide, phosphorothioate, and phosphorodithioate linkages. In other embodiments, the double stranded RNA contains one or two capped strands, as disclosed, for example, by WO 00/63364, filed April 19, 2000 or U.S.S.N. 60/130,377, filed April 21, 1999. In other embodiments, the double stranded RNA contains coding sequence or non-coding sequence, for example, a regulatory sequence (*e.g.*, a transcription factor binding site, a promoter, or a 5' or 3' untranslated region (UTR) of an mRNA). Additionally, the double stranded RNA can be any of the at least partially double-stranded RNA molecules disclosed in WO 00/63364, filed April 19, 2000 (see, for example, pages 8-22). Any of the double stranded RNAs may be expressed *in vitro* or *in vivo* using the methods described herein or standard methods, such as those described in WO 00/63364, filed April 19, 2000 (see, for example, pages 16-22).

[0052] By "double stranded RNA expression library"

or "dsRNA expression library" is meant a collection of nucleic acid expression vectors containing nucleic acid sequences, for example, cDNA sequences or randomized nucleic acid sequences that are capable of forming a double stranded RNA (dsRNA) upon expression of the nucleic acid sequence. Desirably the double stranded RNA expression library contains at least 10,000 unique nucleic acid sequences, more desirably at least 50,000; 100,000; or 500,000 unique nucleic acid sequences, and most desirably, at least 1,000,000 unique nucleic acid sequences. By a "unique nucleic acid sequence" is meant that a nucleic acid sequence of a double stranded RNA expression library has desirably less than 50%, more desirably less than 25% or 20%, and most desirably less than 10% nucleic acid identity to another nucleic acid sequence of a double stranded RNA expression library when the full length sequence are compared. Sequence identity is typically measured using sequence analysis software with the default parameters specified therein (*e.g.*, Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705). This software program matches similar sequences by assigning degrees of homology to various substitutions, deletions, and other modifications.

[0053] The preparation of cDNAs for the generation of double stranded RNA expression libraries is described herein. A randomized nucleic acid library may also be generated as described in detail below. The double stranded RNA expression library may contain nucleic acid sequences that are transcribed in the nucleus or that are transcribed in the cytoplasm of the cell. A double stranded RNA expression library may be generated using techniques described herein.

[0054] By "target nucleic acid" is meant a nucleic acid sequence whose expression is modulated as a result of post-transcriptional gene silencing. As used herein, the target nucleic acid may be in the cell in which the PTGS event occurs or it may be in a neighboring cell, or in a cell contacted with media or other extracellular fluid in which the cell that has undergone the PTGS event is contained. Exemplary target nucleic acids include nucleic acids associated with cancer or abnormal cell growth, such as oncogenes, and nucleic acids associated with an autosomal dominant or recessive disorder. Desirably, the double stranded RNA inhibits the expression of an allele of a nucleic acid that has a mutation associated with a dominant disorder and does not substantially inhibit the other allele of the nucleic acid (*e.g*, an allele without a mutation associated with the disorder). Other exemplary target nucleic acids include host cellular nucleic acids or pathogen nucleic acids required for the infection or propagation of a pathogen, such as a virus, bacteria, yeast, protozoa, or parasite.

[0055] By "target polypeptide" is meant a polypeptide whose biological activity is modulated as a result of post-transcriptional gene silencing. As used herein, the target polypeptide may be in the cell in which the PTGS event occurs or it may be in a neighboring cell, or in a cell contacted with media or other extracellular fluid in which the cell that has undergone the PTGS event is contained.

[0056] As used herein, by "randomized nucleic acids" is meant nucleic acids, for example, those that are at least 100, 500, 600, or 1000 nucleotides in length, constructed from RNA isolated from a particular cell type. In other embodiments, the nucleic acids are at least 10, 20, 30, 40, 50, 60, 70, 80, or 90 nucleotides in length. In yet other embodiments, the number of nucleotides in the nucleic acids is between 5-100 nucleotides, 15-100 nucleotides, 20-95 nucleotides, 25-90 nucleotides, 35-85 nucleotides, 45-80 nucleotides, 50-75 nucleotides, or 55-70 nucleotides, inclusive. In still other embodiments, the number of nucleotides in the nucleic acids is contained in one of the following ranges: 5-15 nucleotides, 15-20 nucleotides, 20-25 nucleotides, 25-35 nucleotides, 35-45 nucleotides, 45-60 nucleotides, 60-70 nucleotides, 70-80 nucleotides, 80-90 nucleotides, or 90-100 nucleotides, inclusive. In other embodiments, the nucleic acids contain less than 50,000; 10,000; 5,000; or 2,000 nucleotides. A randomized nucleic acid library may be constructed in a number of ways. For example, it may be constructed from existing cDNA libraries. In one example, the cDNA libraries are shuffled using the "Gene Shuffling" technology of Maxygen Corp. The cDNA sequences are amplified using inefficient PCR either by restricting elongation time or through the use of manganese. A library of recombinants is created, and the library is finally amplified by PCR and cloned into vectors. In a second method, existing cDNA libraries are digested with an endonuclease to generate fragments of 10 to 300 base pairs. Alternatively, the cDNA libraries are digested to generate shorter fragments of, for example, 5 to 50 base pairs, 5 to 40 base pairs, 5 to 20 base pairs, 5 to 10 base pairs, or 10 to 20 base pairs, inclusive. If the fragments are to contain 5' OH and 3' PO_4 groups, they are dephosphorylated using alkaline phosphatase and phosphorylated using polynucleotide kinase. These dsDNA fragments are then ligated to form larger molecules, and are size selected. In a third example, randomized nucleic acid libraries are created by using random priming of cDNA libraries (using random hexamers and Klenow) to generate short fragments of 20 to 100 nucleotides. Alternatively, shorter fragments are generated that contain, for example, 5 to 50 nucleotides, 5 to 40 nucleotides, 5 to 20 nucleotides, 5 to 10 nucleotides, or 10 to 20 nucleotides, inclusive. These fragments are then ligated randomly to give a desired sized larger fragment.

[0057] Alternatively, a randomized nucleic acid library can be generated from random sequences of oligonucleotides. For example, DNA or RNA oligonucleotides may be prepared chemically. Random DNA sequences may also be prepared enzymatically using terminal transferase in the presence of all dNTPs. Random RNA molecules may be prepared using NDPs and NDP phos-

phorylase. The random sequences may be 10 to 300 bases in length. Alternatively, shorter random sequences are used that contain, for example, 5 to 50 bases, 5 to 40 bases, 5 to 20 bases, 5 to 10 bases, or 10 to 20 bases, inclusive. The sequences are ligated to form the desired larger sequence using RNA ligase. Alternatively these sequences may be ligated chemically. The oligonucleotides are phosphorylated at the 5' position using polynucleotide kinase or by chemical methods, prior to ligation enzymatically. Chemical ligations can utilize a 5' PO_4 and a 3' OH group or a5' OH and a 3' PO_4 group.

[0058] Alternatively, a randomized nucleic acid library can be generated by converting the random DNA sequences into dsDNA sequences using DNA polymerase (Klenow), dNTP and random heteromeric primers, and the RNA sequences are converted into dsDNA sequences by reverse transcriptase and Klenow. After converting into DNA (ss or ds) the sequences are then amplified by PCR. The dsDNA fragments can also be ligated to give larger fragments of a desired size.

[0059] The randomized nucleic acids may be cloned into a vector, for example, an expression vector, as a double stranded RNA transcription cassette. The sequence of the nucleic acid may not be known at the time the vector is generated. The randomized nucleic acid may contain coding sequence or non-coding sequence, for example, a regulatory sequence (*e.g.*, a transcription factor binding site, a promoter, or a 5' or 3' untranslated region (UTR)of an mRNA).

[0060] By "Cre-mediated double recombination" is meant two nucleic acid recombination events involving *loxP* sites that are mediated by Cre recombinase. A Cre-mediated double recombination event can occur, for example, as illustrated in Fig. 1.

[0061] By "function of a cell" is meant any cell activity that can be measured or assessed. Examples of cell function include, but are not limited to, cell motility, apoptosis, cell growth, cell invasion, vascularization, cell cycle events, cell differentiation, cell dedifferentiation, neuronal cell regeneration, and the ability of a cell to support viral replication. The function of a cell may also be to affect the function, gene expression, or the polypeptide biological activity of another cell, for example, a neighboring cell, a cell that is contacted with the cell, or a cell that is contacted with media or other extracellular fluid that the cell is contained in.

[0062] By "apoptosis" is meant a cell death pathway wherein a dying cell displays a set of well-characterized biochemical hallmarks that include cytolemmal membrane blebbing, cell soma shrinkage, chromatin condensation, nuclear disintegration, and DNA laddering. There are many well-known assays for determining the apoptotic state of a cell, including, and not limited to: reduction of MTT tetrazolium dye, TUNEL staining, Annexin V staining, propidium iodide staining, DNA laddering, PARP cleavage, caspase activation, and assessment of cellular and nuclear morphology. Any of these or other known assays may be used in the methods of the invention to determine whether a cell is undergoing apoptosis.

[0063] By "polypeptide biological activity" is meant the ability of a target polypeptide to modulate cell function. The level of polypeptide biological activity may be directly measured using standard assays known in the art. For example, the relative level of polypeptide biological activity may be assessed by measuring the level of the mRNA that encodes the target polypeptide (*e.g.*, by reverse transcription-polymerase chain reaction (RT-PCR) amplification or Northern blot analysis); the level of target polypeptide (*e.g.*, by ELISA or Western blot analysis); the activity of a reporter gene under the transcriptional regulation of a target polypeptide transcriptional regulatory region (*e.g.*, by reporter gene assay, as described below); the specific interaction of a target polypeptide with another molecule, for example, a polypeptide that is activated by the target polypeptide or that inhibits the target polypeptide activity (*e.g.*, by the two-hybrid assay); or the phosphorylation or glycosylation state of the target polypeptide. A compound, such as a dsRNA, that increases the level of the target polypeptide, mRNA encoding the target polypeptide, or reporter gene activity within a cell, a cell extract, or other experimental sample is a compound that stimulates or increases the biological activity of a target polypeptide. A compound, such as a dsRNA, that decreases the level of the target polypeptide, mRNA encoding the target polypeptide, or reporter gene activity within a cell, a cell extract, or other experimental sample is a compound that decreases the biological activity of a target polypeptide.

[0064] By "assaying" is meant analyzing the effect of a treatment, be it chemical or physical, administered to whole animals, cells, tissues, or molecules derived therefrom. The material being analyzed may be an animal, a cell, a tissue, a lysate or extract derived from a cell, or a molecule derived from a cell. The analysis may be, for example, for the purpose of detecting altered cell function, altered gene expression, altered endogenous RNA stability, altered polypeptide stability, altered polypeptide levels, or altered polypeptide biological activity. The means for analyzing may include, for example, antibody labeling, immunoprecipitation, phosphorylation assays, glycosylation assays, and methods known to those skilled in the art for detecting nucleic acids. In some embodiments, assaying is conducted under selective conditions.

[0065] By "modulates" is meant changing, either by a decrease or an increase. As used herein, desirably a nucleic acid sequence decreases the function of a cell, the expression of a target nucleic acid in a cell, or the biological activity of a target polypeptide in a cell by least 20%, more desirably by at least 30%, 40%, 50%, 60% or 75%, and most desirably by at least 90%. Also as used herein, desirably a nucleic acid sequence increases the function of a cell, the expression of a target nucleic acid in a cell, or the biological activity of a target

polypeptide in a cell by at least 1.5-fold to 2-fold, more desirably by at least 3-fold, and most desirably by at least 5-fold.

[0066] By "a decrease" is meant a lowering in the level of: a) protein (*e.g.*, as measured by ELISA or Western blot analysis); b) reporter gene activity (*e.g.*, as measured by reporter gene assay, for example, β-galactosidase, green fluorescent protein, or luciferase activity); c) mRNA (*e.g.*, as measured by RT-PCR or Northern blot analysis relative to an internal control, such as a "housekeeping" gene product, for example, β-actin or glyceraldehyde 3-phosphate dehydrogenase (GAPDH)); or d) cell function, for example, as assayed by the number of apoptotic, mobile, growing, cell cycle arrested, invasive, differentiated, or dedifferentiated cells in a test sample. In all cases, the lowering is desirably by at least 20%, more desirably by at least 30%, 40%, 50%, 60%, 75%, and most desirably by at least 90%. As used herein, a decrease may be the direct or indirect result of PTGS.

[0067] By "an increase" is meant a rise in the level of: a) protein (*e.g.*, as measured by ELISA or Western blot analysis); b) reporter gene activity (*e.g.*, as measured by reporter gene assay, for example, β-galactosidase, green fluorescent protein, or luciferase activity); c) mRNA (*e.g.*, as measured by RT-PCR or Northern blot analysis relative to an internal control, such as a "housekeeping" gene product, for example, β-actin or glyceraldehyde 3-phosphate dehydrogenase (GAPDH)); or d) cell function, for example, as assayed by the number of apoptotic, mobile, growing, cell cycle arrested, invasive, differentiated, or dedifferentiated cells in a test sample. Desirably, the increase is by at least 1.5-fold to 2-fold, more desirably by at least 3-fold, and most desirably by at least 5-fold. As used herein, an increase may be the indirect result of PTGS. For example, the double stranded RNA may inhibit the expression of a protein, such as a suppressor protein, that would otherwise inhibit the expression of another nucleic acid.

[0068] By "alteration in the level of gene expression" is meant a change in transcription, translation, or mRNA or protein stability such that the overall amount of a product of the gene, *i.e.*, mRNA or polypeptide, is increased or decreased.

[0069] By "reporter gene" is meant any gene that encodes a product whose expression is detectable and/or able to be quantitated by immunological, chemical, biochemical, or biological assays. A reporter gene product may, for example, have one of the following attributes, without restriction: fluorescence (*e.g.*, green fluorescent protein), enzymatic activity (*e.g.*, β-galactosidase, luciferase, chloramphenicol acetyltransferase), toxicity (*e.g.*, ricin A), or an ability to be specifically bound by an additional molecule (*e.g.*, an unlabeled antibody, followed by a labelled secondary antibody, or biotin, or a detectably labelled antibody). It is understood that any engineered variants of reporter genes that are readily available to one skilled in the art, are also included, without restriction, in the foregoing definition.

[0070] By "protein" or "polypeptide" or "polypeptide fragment" is meant any chain of more than two amino acids, regardless of post-translational modification (*e.g.*, glycosylation or phosphorylation), constituting all or part of a naturally-occurring polypeptide or peptide, or constituting a non-naturally occurring polypeptide or peptide.

[0071] By "promoter" is meant a minimal sequence sufficient to direct transcription of a gene. Also included in this definition are those transcription control elements (*e.g.*, enhancers) that are sufficient to render promoter-dependent gene expression controllable in a cell type-specific, tissue-specific, or temporal-specific manner, or that are inducible by external signals or agents; such elements, which are well-known to skilled artisans, may be found in a 5' or 3' region of a gene or within an intron. Desirably a promoter is operably linked to a nucleic acid sequence, for example, a cDNA or a gene in such a way as to permit expression of the nucleic acid sequence.

[0072] By "operably linked" is meant that a gene and one or more transcriptional regulatory sequences, *e.g.*, a promoter or enhancer, are connected in such a way as to permit gene expression when the appropriate molecules (*e.g.*, transcriptional activator proteins) are bound to the regulatory sequences.

[0073] By "expression vector" is meant a DNA construct that contains at least one promoter operably linked to a downstream gene or coding region (*e.g.*, a cDNA or genomic DNA fragment that encodes a protein, optionally, operatively linked to sequence lying outside a coding region, an antisense RNA coding region, or RNA sequences lying outside a coding region). Transfection or transformation of the expression vector into a recipient cell allows the cell to express RNA encoded by the expression vector. An expression vector may be a genetically engineered plasmid, virus, or artificial chromosome derived from, for example, a bacteriophage, adenovirus, retrovirus, poxvirus, or herpesvirus.

[0074] By "transformation" or "transfection" is meant any method for introducing foreign molecules into a cell (*e.g.*, a bacterial, yeast, fungal, algal, plant, insect, or animal cell, particularly a vertebrate or mammalian cell). The cell may be in an animal. Lipofection, DEAE-dextran-mediated transfection, microinjection, protoplast fusion, calcium phosphate precipitation, viral or retroviral delivery, electroporation, and biolistic transformation are just a few of the transformation/transfection methods known to those skilled in the art. The RNA or RNA expression vector (DNA) may be naked RNA or DNA or local anesthetic complexed RNA or DNA (Pachuk *et al.*, *supra*). Other standard transformation/transfection methods and other RNA and/or DNA delivery agents (*e.g.*, a cationic lipid, liposome, or bupivacaine) are described in WO 00/63364, filed April 19, 2000 (see, for example, pages 18-26). Commercially available kits can also be used to deliver RNA or DNA to a cell. For example, the Transmessenger Kit from Qiagen, an RNA kit

from Xeragon Inc., and an RNA kit from DNA Engine Inc. (Seattle, WA) can be used to introduce single or double stranded RNA into a cell.

[0075] By "transformed cell" or "transfected cell" is meant a cell (or a descendent of a cell) into which a nucleic acid molecule, for example, a double stranded RNA or double stranded expression vector has been introduced, by means of recombinant nucleic acid techniques. Such cells may be either stably or transiently transfected.

[0076] By "selective conditions" is meant conditions under which a specific cell or group of cells can be selected for. For example, the parameters of a fluorescence-activated cell sorter (FACS) can be modulated to identify a specific cell or group of cells. Cell panning, a technique known to those skilled in the art, is another method that employs selective conditions.

[0077] As use herein, by "optimized" is meant that a nucleic acid fragment is generated through inefficient first strand synthesis (*e.g.*, reverse transcription (RT) and/or RT/second strand synthesis (RT-SSS) using Klenow or other enzymes and/or RT-PCR or PCR, to be of a particular length. Desirably the length of the nucleic acid fragment is less than a full length cDNA or is 100, 500, 600, or 1000 nucleotides in length. In other embodiments, the nucleic acid fragment is at least 10, 20, 30, 40, 50, 60, 70, 80, or 90 nucleotides in length. In yet other embodiments, the number of nucleotides in the nucleic acid fragment is between 5-100 nucleotides, 15-100 nucleotides, 20-95 nucleotides, 25-90 nucleotides, 35-85 nucleotides, 45-80 nucleotides, 50-75 nucleotides, or 55-70 nucleotides, inclusive. In still other embodiments, the number of nucleotides in the nucleic acid fragment is contained in one of the following ranges: 5-15 nucleotides, 15-20 nucleotides, 20-25 nucleotides, 25-35 nucleotides, 35-45 nucleotides, 45-60 nucleotides, 60-70 nucleotides, 70-80 nucleotides, 80-90 nucleotides, or 90-100 nucleotides, inclusive. In other embodiments, the nucleic acid fragment contains less than 50,000; 10,000; 5,000; or 2,000 nucleotides. Optimization of the length of a nucleic acid can be achieved during first strand or second strand synthesis of a desired nucleic acid by lowering Mg^{++} concentrations to no less than the nucleotide concentrations; by adding Mn^{++} to the reaction to achieve the desired size selection (*e.g.*, by replacing Mg^{++} completely, or by adding Mn^{++} at varying concentrations along with Mg^{++}); by decreasing and/or limiting concentrations of dNTP(s) to effect the desired fragment size; by using various concentrations of ddNTP(s) along with standard or optimal concentrations of dNTP(s), to achieve varying ratios, to obtain the desired fragment size; by using limited and controlled exonuclease digestion of the fragment following RT, RT-SSS, RT-PCR, or PCR; or by a combination of any of these methods.

[0078] As used herein, by "sized selected" is meant that a nucleic acid of a particular size is selected for use in the construction of dsRNA expression libraries as described herein. Desirably the size selected nucleic acid is less than a full length cDNA sequence or at least 100, 500, 600, or 1000 nucleotides in length. In other embodiments, the nucleic acid is at least 10, 20, 30, 40, 50, 60, 70, 80, or 90 nucleotides in length. In yet other embodiments, the number of nucleotides in the nucleic acid is between 5-100 nucleotides, 15-100 nucleotides, 20-95 nucleotides, 25-90 nucleotides, 35-85 nucleotides, 45-80 nucleotides, 50-75 nucleotides, or 55-70 nucleotides, inclusive. In still other embodiments, the number of nucleotides in the nucleic acid is contained in one of the following ranges: 5-15 nucleotides, 15-20 nucleotides, 20-25 nucleotides, 25-35 nucleotides, 35-45 nucleotides, 45-60 nucleotides, 60-70 nucleotides, 70-80 nucleotides, 80-90 nucleotides, or 90-100 nucleotides, inclusive. In other embodiments, the nucleic acid contains less than 50,000; 10,000; 5,000; or 2,000 nucleotides. For example, a nucleic acid may be size selected using size exclusion chromatography (*e.g.*, as size exclusion Sepharose matrices) according to standard procedures (see, for example, Sambrook, Fritsch, and Maniatis, Molecular Cloning: A Laboratory Manual (2d ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

[0079] By "under conditions that inhibit or prevent an interferon response or a dsRNA stress response" is meant conditions that prevent or inhibit one or more interferon responses or cellular RNA stress responses involving cell toxicity, cell death, an anti-proliferative response, or a decreased ability of a dsRNA to carry out a PTGS event. These responses include, but are not limited to, interferon induction (both Type 1 and Type II), induction of one or more interferon stimulated genes, PKR activation, 2'5'-OAS activation, and any downstream cellular and/or organismal sequelae that result from the activation/induction of one or more of these responses. By "organismal sequelae" is meant any effect (s) in a whole animal, organ, or more locally (*e.g.*, at a site of injection) caused by the stress response. Exemplary manifestations include elevated cytokine production, local inflammation, and necrosis. Desirably the conditions that inhibit these responses are such that not more than 95%, 90%, 80%, 75%, 60%, 40%, or 25%, and most desirably not more than 10% of the cells undergo cell toxicity, cell death, or a decreased ability to carry out a PTGS event, compared to a cell not exposed to such interferon response inhibiting conditions, all other conditions being equal (*e.g.*, same cell type, same transformation with the same dsRNA expression library.

[0080] Apoptosis, interferon induction, 2'5' OAS activation/induction, PKR induction/activation, anti-proliferative responses, and cytopathic effects are all indicators for the RNA stress response pathway. Exemplary assays that can be used to measure the induction of an RNA stress response as described herein include a TUNEL assay to detect apoptotic cells, ELISA assays to detect the induction of alpha, beta and gamma interferon, ribosomal RNA fragmentation analysis to detect

activation of 2'5'OAS, measurement of phosphorylated eIF2a as an indicator of PKR (protein kinase RNA inducible) activation, proliferation assays to detect changes in cellular proliferation, and microscopic analysis of cells to identify cellular cytopathic effects (see, *e.g.*, Example 11). Desirably, the level of an interferon response or a dsRNA stress response in a cell transformed with a double stranded RNA or a double straned RNA expression vector is less than 20, 10, 5,or 2-fold greater than the corresponding level in a mock-transfected control cell under the same conditions, as measured using one of the assays described herein. In other embodiments, the level of an interferon response or a dsRNA stress response in a cell transformed with a double stranded RNA or a double straned RNA expression vector using the methods of the present invention is less than 500%, 200%, 100%, 50%, 25%, or 10% greater than the corresponding level in a corresponding transformed cell that is not exposed to such interferon response inhibiting conditions, all other conditions being equal. Desirably, the double straned RNA does not induce a global inhibition of cellular transcription or translation.

[0081] By □specifically hybridizes □is meant a double stranded RNA that hybridizes to a target nucleic acid but does not substantially hybridize to other nucleic acids in a sample (*e.g.*, a sample from a cell) that naturally includes the target nucleic acid, when assayed under denaturing conditions. In one embodiment, the amount of a target nucleic acidhybridized to, or associated with, the double stranded RNA, as measured using standard assays, is 2-fold, desirably 5-fold, more desirably 10-fold, and most desirably 50-fold greater than the amount of a control nucleic acid hybridized to, or associated with, the double stranded RNA.

[0082] By □high stringency conditions □is meant hybridization in 2X SSC at 40°C with a DNA probe length of at least 40 nucleotides. For other definitions of high stringency conditions, see F. Ausubel *et al., Current Protocols in Molecular Biology*, pp. 6.3.1-6.3.6, John Wiley & Sons, New York, NY, 1994, hereby incorporated by reference.

[0083] Conditions and techniques that can be used to prevent an interferon response or dsRNA stress response during the screening methods of the present invention are described herein.

Brief Description of the Drawing

[0084] Fig. 1 is a schematic representation of a strategy to isolate clonally pure stable integrants that contain a single expression unit isolated from cells transfected with a double-stranded RNA encoding a cDNA library.

[0085] Fig. 2 is a schematic illustration of the production of effector RNAs in cells expressing PSA. The PSA expression cassette used to create the transient PSA expression cell line is depicted at the top of the figure. Expression of PSA is driven by the HCMV IE promoter and the SV40 polyadenylation signal (pA). Only sequences 3' of the PSA initiation codon have been used in these vectors. The effector RNA expression cassettes are shown below the PSA expression cassette and are designed to express PSA sense RNA, PSA antisense RNA, and PSA dsRNA. Expression of the effector RNAs is under the control of the T7 promoter (T7p). Transcription from T7p is catalyzed by T7 RNA polymerase, which is supplied by co-transfecting a T7 RNA polymerase expression plasmid (not shown). Control effector RNA cassettes expressing irrelevant RNAs derived from the Herpes simplex virus glycoprotein D gene were included as controls. The 600 base pair sequence from the Herpes simplex gD gene is from Herpes Simplex virus 2 strain 12 and maps to the coding region downstream of the gD initiation codon.

[0086] Fig. 3 is a bar graph illustrating silencing of PSA expression by dsRNA. PSA levels in the supernates of transfected cells were determined by ELISA and are plotted as percent expression of the PSA untreated control. The PSA untreated control shown at the left is normalized to 100%. PSA levels in the supernates of cells transfected with the various effector PSA or control RNAs are shown by the shaded and open bars respectively. All data shown is from day two post-transfection. Data from later time points were similar to the day two time point.

[0087] Fig. 4 is a schematic illustration of the RNA stress response pathway, also known as the Type 1 interferon response.

Detailed Description of the Invention

[0088] Post-transcriptional gene silencing (PTGS) can be used as a tool to identify and validate specific unknown genes involved in cell function, gene expression, and polypeptide biological activity. Although the use of PTGS as a validation strategy is well documented in invertebrates and plants, its use in identification of genes that modulate cell function, gene expression, or polypeptide biological activity, as described below, is novel. Since novel genes are likely to be identified through the methods of the present invention, PTGS is developed for use in validation and to identify novel targets for use in theraples for diseases, for example, cancer, neurological disorders, obesity, leukemia, lymphomas, and other disorders of the blood or immune system.

[0089] The present invention features methods to identify unknown targets that result in the modulation of a particular phenotype, an alteration of gene expression in a cell, or an alteration in polypeptide biological activity in a cell, using either a library based screening approach or a non-library based approach to identify nucleic acids that induce gene silencing. The present invention also allows the determination of function of a given sequence. These methods involve the direct delivery of *in vitro* transcribed double stranded RNA (dsRNA), as well

as plasmid-based systems that direct the cell to make its own dsRNA. To avoid problems associated with transfection efficiency, plasmids are designed to contain a selectable marker to ensure the survival of only those cells that have taken up plasmid DNA. One group of plasmids directs the synthesis of dsRNA that is transcribed in the cytoplasm, while another group directs the synthesis of dsRNA that is transcribed in the nucleus.

Identification of genes by assaying for a modulation in cell function

[0090] Functional identification of novel genes can be accomplished through the use of a number of different assays. For example, cells may be assayed for cell motility, apoptosis, cell growth, cell invasion, vascularization, cell cycle events, cell differentiation, cell dedifferentiation, neuronal cell regeneration, or the ability to support viral replication, as well as other cell functions known in the art. Methods for carrying out such functional assays are well known and are described, for example, in Platet and Garcia (Invasion Metastasis 18: 198-208, 1998-1999); Harper *et al.* (Neuroscience 88: 257-267, 1999); and Tomaselli *et al.* (J. Cell Biol. 105: 2347-2358, 1987), and are also described below.

[0091] Functional identification of nucleic acid sequences involved in modulating a particular cell function may be carried out by comparing cells transfected with a dsRNA to control cells that have not been transformed with a dsRNA or that have been mock-transfected, in a functional assay. A cell that has taken up sequences unrelated to a particular function will perform in the particular assay in a manner similar to the control cell. A cell experiencing PTGS of a gene involved in the particular function will exhibit an altered ability to perform in the functional assay compared to the control.

[0092] The percent modulation of a particular cell function that identifies a nucleic acid sequence that modulates the function of a cell will vary depending on the assay, phenotype, and the particular nucleic acid affected by PTGS. For each assay, the percent modulation can readily be determined by one skilled in the art, when used in conjunction with controls, as described herein. Desirably the modulation is at least 20%, more desirably at least 30%, 40%, 50%, 60%, 75%, and most desirably at least 90% compared to the control. An increase in the function of a cell can also be measured in terms of fold increase, where desirably, the increase is at least 1.5-fold to 5-fold compared to the control.

[0093] Alternatively, the function of a cell may be to affect the function, gene expression, or polypeptide biological activity of another cell, for example, a neighboring cell, a cell that is contacted with the cell in which a PTGS event occurs, or a cell that is contacted with media or other extracellular fluid that the cell in which a PTGS event occurs is contained in. For example, a cell experiencing PTGS of a gene may modulate cell motility, apoptosis, cell growth, cell invasion, vascularization, cell cycle events, cell differentiation, cell dedifferentiation, neuronal cell regeneration, or the ability to support viral replication of a nearby cell, or a cell that is exposed to media or other extracellular fluid in which the transfected cell in which a PTGS event occurs was once contained. This can be tested by removing the media in which a cell experiencing a PTGS event is occurring and placing it on a separate cell or population of cells. If the function of the separate cell or population of cells is modulated, compared to a cell or population of cells receiving media obtained from cells that had been mock transfected, then one or more of the cells experiencing a PTGS event can affect the function of another cell. The identity of the nucleic acid sequence that causes the modulation can be identified with repeated rounds of selection.

[0094] In another method, a single cell experiencing a PTGS event can be placed in proximity of a cell or a population of cells that was not transfected with dsRNA, and the effect of this placement is evaluated for a modulation in the function of the cell or population of cells. If the function of the non-transfected cell or population of cells is modulated, compared to a cell or population of cells in proximity of a cell that was mock transfected, then the cell experiencing a PTGS event contains a nucleic acid sequence that can affect the function of another cell. This nucleic acid sequence can be identified using techniques described herein.

Identification of genes using differential gene expression

[0095] Differential gene expression analysis can be used to identify a nucleic acid sequence that modulates the expression of a target nucleic acid in a cell. Alterations in gene expression induced by gene silencing can be monitored in a cell into which a dsRNA has been introduced. For example, differential gene expression can be assayed by comparing nucleic acids expressed in cells into which dsRNA has been introduced to nucleic acids expressed in control cells that were not transfected with dsRNA or that were mock-transfected. Gene array technology can be used in order to simultaneously examine the expression levels of many different nucleic acids. Examples of methods for such expression analysis are described by Marrack *et al.* (Current Opinions in Immunology 12:206-209, 2000); Harkin (Oncologist 5:501-507, 2000); Pelizzari *et al.* (Nucleic Acids Res. 28: 4577-4581, 2000); and Marx (Science 289:1670-1672, 2000).

Identification of genes by assaying polypeptide biological activity

[0096] Novel nucleic acid sequences that modulate the biological activity of a target polypeptide can also be identified by examining polypeptide biological activity. Various polypeptide biological activities can be evaluat-

ed to identify novel genes according to the methods of the invention. For example, the expression of a target polypeptide(s) may be examined. Alternatively, the interaction between a target polypeptide(s) and another molecule(s), for example, another polypeptide or a nucleic acid may be assayed. Phosphorylation or glycosylation of a target polypeptide(s) may also be assessed, using standard methods known to those skilled in the art.

[0097] Identification of nucleic acid sequences involved in modulating the biological activity of a target polypeptide may be carried out by comparing the polypeptide biological activity of a cell transfected with a dsRNA to a control cell that has not been transfected with a dsRNA or that has been mock-transfected. A cell that has taken up sequences unrelated to a particular polypeptide biological activity will perform in the particular assay in a manner similar to the control cell. A cell experiencing PTGS of a gene involved in the particular polypeptide biological activity will exhibit an altered ability to perform in the biological assay, compared to the control.

Insertion of single units into the chromosome and generation of a cell line containing a single dsRNA expression library integrant

[0098] The present invention involves the generation of a target cell line in which the dsRNA expression library is subsequently introduced. Through the use of site-specific recombination, single integrants of dsRNA expression cassettes are generated at the same locus of all cells in the target cell line, allowing uniform expression of the dsRNA in all of the integrants. A dsRNA expression library derived from various cell lines is used to create a representative library of stably integrated cells, each cell within the target cell line containing a single integrant. Cre/*lox*, Lambda-Cro repressor, and Flp recombinase systems or retroviruses are used to generate these singular integrants of dsRNA expression cassettes in the target cell line (Satoh *et al.*, J. Virol. 74: 10631-10638, 2000; Trinh *et al.*, J. Immunol. Methods 244:185-193, 2000; Serov *et al.*, An. Acad. Bras. Cienc. 72:389-398, 2000; Grez *et al.*, Stem Cells. 16:235-243, 1998; Habu *et al.*, Nucleic Acids Symp. Ser. 42:295-296, 1999; Haren *et al.*, Annu. Rev. Microbiol. 53:245-281, 1999; Baer *et al.*, Biochemistry 39:7041-7049, 2000; Follenzi *et al.*, Nat. Genet. 25:217-222, 2000; Hindmarsh *et al.*, Microbiol. Mol. Biol. Rev. 63:836-843, 1999; Darquet *et al.*, Gene Ther. 6:209-218, 1999; Darquet *et al.*, Gene Ther. 6:209-218, 1999; Yu *et al.*, Gene 223:77-81, 1998; Darquet *et al.*, Gene Ther. 4: 1341-1349, 1997; and Koch *et al.*, Gene 249:135-144, 2000). These systems are used singularly to generate singular insertion clones, and also in combination.

[0099] The following exemplary sequence specific integrative systems use short target sequences that allow targeted recombination to be achieved using specific proteins: FLP recombinase, bacteriophage Lambda integrase, HIV integrase, and pilin recombinase of Salmonella (Seng *et al.* Construction of a Flp "exchange cassette" contained vector and gene targeting in mouse ES cell] A book chapter PUBMED entry 11797223 - Sheng Wu Gong Cheng Xue Bao. 2001 Sep;17(5):566-9., Liu *et al.*, Nat Genet. 2001 Jan 1;30(1):66-72., Awatramani *et al.*, Nat Genet. 2001 Nov;29(3):257-9., Heidmann and Lehner, Dev Genes Evol. 2001 Sep;211(8-9):458-65, Schaft *et al.*, Genesis. 2001 Sep;31(1):6-10, Van Duyne, Annu Rev Biophys Biomol Struct. 2001;30: 87-104., Lorbach *et al.*, J Mol Biol. 2000 Mar 10;296(5): 1175-81., Darquet *et al.*, Gene Ther. 1999 Feb;6(2): 209-18., Bushman and Miller, J Virol. 1997 Jan;71(1): 458-64., Fulks *et al.*, J Bacteriol. 1990 Jan;172(1): 310-6). A singular integrant is produced by randomly inserting the specific sequence (*e.g.*, loxP in the cre recombinase system) and selecting or identifying the cell that contains a singular integrant that supports maximal expression. For example, integrants that show maximal expression following random integration can be identified through the use of reporter gene sequences associated with the integrated sequence. The cell can be used to specifically insert the expression cassette into the site that contains the target sequence using the specific recombinase, and possibly also remove the expression cassette that was originally placed to identify the maximally expressing chromosomal location. A skilled artisan can also produce singular integrants using retroviral vectors, which integrate randomly and singularly into the eukaryotic genome. In particular, singular integrants can be produced by inserting retroviral vectors that have been engineered to contain the desired expression cassette into a naive cell and selecting for the chromosomal location that results in maximal expression (Michael *et al.*, EMBO Journal, vol 20: pages 2224-2235, 2001; Reik and Murrell., Nature, vol. 405, page 408-409, 2000; Berger *et al.*, Molecular Cell, vol. 8, pages 263-268). One may also produce a singular integrant by cotransfecting the bacterial RecA protein with or without nuclear localization signal along with sequences that are homologous to the target sequence (*e.g.*, a target endogenous sequence or integrated transgene sequence). Alternatively, a nucleic acid sequence that encodes a RecA protein with nuclear localization signals can be cotransfected (Shibata *et al.*, Proc Natl Acad Sci U S A. 2001 Jul 17;98(15):8425-32. Review., Muyrers *et al.*, Trends Biochem Sci. 2001 May;26(5): 325-31., Paul *et al.*, Mutat Res. 2001 Jun 5;486(1):11-9., Shcherbakova *et al.*, Mutat Res. 2000 Feb 16;459(1): 65-71., Lantsov. Mol Biol (Mosk). 1994 May-Jun;28(3): 485-95).

[0100] An example utilizing such methods is detailed below.

Creation of the target cell line

[0101] Target cell lines are the same cell lines as the

ones from which the dsRNA expression libraries will be derived. Target cells are created by transfecting the selected cell line with a bicistronic plasmid expressing a selectable marker, such as G418 and the reporter gene GFP. The plasmid also bears a *loxP* site. Plasmids integrate randomly into the chromosome through the process of illegitimate recombination at a frequency of 10^{-4}. Following transfection, cells containing integrants are selected by culturing the cells in the presence of G418 at a concentration determined earlier in a kill curve analysis. About a dozen G418-resistant colonies are expanded and relative GFP expression levels are determined using flow cytometry. DNA from the cells is analyzed by Southern blot analysis to determine integrant copy number. Several single copy integrants exhibiting the highest GFP expression levels are then selected as the target cell lines. GFP expression is monitored because dsRNA encoding templates are then integrated into the loci containing the *loxP*, GFP, and G418 cassettes in a site-specific fashion, and it is important to ensure that these loci are transcriptionally active. Since cells are selected on the basis of G418 resistance and GFP expression, integration of the plasmid DNA can occur at the *loxP* site, destroying its function. Several cell lines are therefore chosen to reasonably ensure that at least one integrant has an intact *loxP* site.

Double stranded RNA expression library construction and site-specific recombination into the target cell line

[0102] A cDNA library or a randomized library is constructed from RNA isolated from selected cell lines. cDNAs or randomized nucleic acids in the size range of at least 100 to 1000 nucleotides, for example, 500 to 600 nucleotides are optimized during synthesis or are size-selected prior to cloning. In other embodiments, the nucleic acids are at least 10, 20, 30, 40, 50, 60, 70, 80, or 90 nucleotides in length. In yet other embodiments, the number of nucleotides in the nucleic acids is between 5-100 nucleotides, 15-100 nucleotides, 20-95 nucleotides, 25-90 nucleotides, 35-85 nucleotides, 45-80 nucleotides, 50-75 nucleotides, or 55-70 nucleotides, inclusive. In still other embodiments, the number of nucleotides in the nucleic acids is contained in one of the following ranges: 5-15 nucleotides, 15-20 nucleotides, 20-25 nucleotides, 25-35 nucleotides, 35-45 nucleotides, 45-60 nucleotides, 60-70 nucleotides, 70-80 nucleotides, 80-90 nucleotides, or 90-100 nucleotides, inclusive. In other embodiments, the nucleic acid contains less than 50,000; 10,000; 5,000; or 2,000 nucleotides. Each cDNA or randomized nucleic acid is then cloned into a plasmid vector as a dsRNA transcription cassette flanked by two convergent promoters (such as T7 promoters as described herein). The promoters are transcriptionally regulated such that they are off until induced, for example, using a tet ON/OFF system (Forster *et al.*, Nucleic Acids Res. 27:7708-710, 1999; Liu *et al.*, Biotechniques 24:624-628, 6,30-632, 1998; and Gatz, Methods Cell Biol. 50:411-424, 1995). The plasmid also contains the hygromycin resistance gene and an inverted *loxP* site. The cDNA plasmid library or randomized plasmid library is then co-transfected into the target cell line with a plasmid expressing Cre recombinase, which catalyzes site-specific recombination of the transfected cDNA plasmid or randomized nucleic acid plasmid at the inverted *loxP* site into the chromosomal locus containing the GFP gene and *loxP* site (see Fig. 1). The use of the Cre/*lox* system allows the efficient integration of a plasmid into the chromosome (every transfected cell is predicted to undergo a plasmid integration event). Other site-specific recombination strategies can also be utilized. This results in having every integration to occur at the same site, thereby obviating potential problems with loci dependent expression.

[0103] Two days following transfection, cells are incubated in the presence of hygromycin to kill untransfected cells and to select for stable integrants. Transcription of dsRNA is induced, and selected cells are assayed for an alteration in cell function, the biological activity of a target polypeptide, or differential gene expression. Cells expressing dsRNA corresponding to a target nucleic acid exhibit an altered function, for example, increased or decreased cell invasion, motility, apoptosis, growth, differentiation, dedifferentiation, or regeneration, or the ability of the cell to support viral replication. Cells exhibiting altered function are then expanded and the sequence of the integrant is determined. Targets are identified and validated using dsRNA specific for the identified target, or other non-PTGS mediated methods, for example antisense technology.

[0104] The regulated transcription system of the present invention provides an advantage to other double stranded expression systems. Following transfection of the dsRNA library, cells contain hundreds to thousands of dsRNA expression cassettes, with concomitant expression of that many expression cassettes. In the dsRNA expression system of the present invention, dsRNA expression cassettes contained within the expression vector integrate into the chromosome of the transfected cell. As described in detail below, every transfected cell integrates one of the double stranded expression cassettes. Desirably no transcription occurs until the episomal (non-integrated) expression vectors are diluted out of the cell such that not more than 5 episomal vectors remain in the cell. Most desirably no transcription occurs until all of the episomal (non-integrated) expression vectors are diluted out of the cell and only the integrated expression cassette remains (a process usually taking about two to several weeks of cell culture). At this time transcription is induced, allowing dsRNA to be expressed in the cells. This method ensures that only one species of dsRNA is expressed per cell, as opposed to other methods that express hundreds to thousands of double stranded species. The use of the above-described system results in the loss of all but one expression cassette, which in turn, permits the rapid

screening of libraries without requiring screening multiple pools of libraries to identify the target gene.

Non-library approaches for the identification of a nucleic acid sequence that modulates cell function, gene expression in a cell, or the biological activity of a target polypeptide in a cell through PTGS

[0105] Nucleic acid sequences that modulate cell function, gene expression in a cell, or the biological activity of a target polypeptide in a cell may also be identified using non-library based approaches involving PTGS. For example, a single known nucleic acid sequence encoding a polypeptide with unknown function or a single nucleic acid fragment of unknown sequence and/or function can be made into a double stranded RNA molecule. This dsRNA is then transfected into a desired cell type and the cell is assayed for modulations in cell function, gene expression of a target nucleic acid in the cell, or the biological activity of a target polypeptide in the cell, using methods described herein. A modulation in cell function, gene expression in the cell, or the biological activity of a target polypeptide in the cell identifies the nucleic acid of the dsRNA as a nucleic acid the modulates the specific cell function, gene expression, or the biological activity of a target polypeptide. As a single dsRNA species is transfected into the cells, the nucleic acid sequence responsible for the modulation is readily identified. This non-library based approach to nucleic acid identification is desirably used under conditions that inhibit an interferon response or dsRNA stress response. Such conditions are described in detail herein.

[0106] The discovery of novel genes through the methods of the present invention may lead to the generation of novel therapeutics. For example, genes that decrease cell invasion may be used as targets for drug development, such as for the development of cytostatic therapeutics for use in the treatment of cancer. Development of such therapeutics is important because currently available cytotoxic anticancer agents are also toxic for normal rapidly dividing cells. In contrast, a cytostatic agent may only need to check metastatic processes, and by inference, slow cell growth, in order to stabilize the disease. In another example, genes that increase neuronal regeneration may be used to develop therapeutics for the treatment, prevention, or control of a number of neurological diseases, including Alzheimer's disease and Parkinson's disease. Genes that are involved in the ability to support viral replication and be used as targets in anti-viral therapies. Such therapies may be used to treat, prevent, or control viral diseases involving human immunodeficiency virus (HIV), hepatitis C virus (HCV), hepatitis B virus (HBV), and human papillomavirus (HPV). The efficacies of therapeutics targeting the genes identified according to the present invention can be further tested in cell culture assays, as well as in animal models.

The use of vertebrate or mammalian cell lines for identification of nucleic acid sequences that modulate cell function, expression of a target nucleic acid or biological activity of a target polypeptide

[0107] While the use of the present invention is not limited to vertebrate or mammalian cells, such cells can be used to carry out the nucleic acid identification methods described herein. Desirably the vertebrate or mammalian cells used to carry out the present invention are cells that have been cultured for only a small number of passages (*e.g.*, less than 30 passages of a cell line that has been obtained directly from American Type Culture Collection), or are primary cells. In addition, vertebrate or mammalian cells can be used to carry out the present invention when the dsRNA being transfected into the cell is not complexed with cationic lipids.

[0108] The following examples are to illustrate the invention. They are not meant to limit the invention in any way. For example, it is noted that any of the following examples can be used with double stranded RNAs of any length. The methods of the present invention can be readily adapted by one skilled in the art to utilize double stranded RNAs of any desired length.

Example 1: Design and delivery of vectors for intracellular synthesis of dsRNA for library based screening approaches to nucleic acid identification using PTGS

[0109] PTGS is induced when dsRNA is made intracellularly. The library based screening approaches to nucleic acid identification through PTGS may require that dsRNA reside in certain cellular compartments in order to exert its effect. Therefore, expression plasmids that transcribe dsRNA in the cytoplasm and in the nucleus are utilized. There are two classes of nuclear transcription vectors: one that is designed to express polyadenylated dsRNA (for example, a vector containing an RNA polymerase II promoter and a poly A site) and one that expresses non-adenylated dsRNA (for example, a vector containing an RNA polymerase II promoter and no poly A site, or a vector containing a T7 promoter). Different cellular distributions are predicted for the two species of RNA; both vectors are transcribed in the nucleus, but the ultimate destinations of the RNA species are different intracellular locations. Intracellular transcription may also utilize bacteriophage T7 and SP6 RNA polymerase, which may be designed to transcribe in the cytoplasm or in the nucleus. Alternatively, Qbeta replicase RNA-dependent RNA polymerase may be used to amplify dsRNA. Viral RNA polymerases, either DNA and RNA dependent, may also be used. Alternatively, dsRNA replicating polymerases can be used. Cellular polymerases such as RNA Polymerase I, II, or III or mitochondrial RNA polymerase may also be utilized. Both the cytoplasmic and nuclear transcription vectors contain an antibiotic resistance gene to enable selection

of cells that have taken up the plasmid. Cloning strategies employ chain reaction cloning (CRC), a one-step method for directional ligation of multiple fragments (Pachuk *et al.*, Gene 243:19-25, 2000). Briefly, the ligations utilize bridge oligonucleotides to align the DNA fragments in a particular order and ligation is catalyzed by a heat-stable DNA ligase, such as Ampligase, available from Epicentre.

Inducible or repressible transcription vectors for the generation of a dsRNA expression library

[0110] If desired, inducible and repressible transcription systems can be used to control the timing of the synthesis of dsRNA. Inducible and repressible regulatory systems involve the use of promoter elements that contain sequences that bind prokaryotic or eukaryotic transcription factors upstream of the sequence encoding dsRNA. In addition, these factors also carry protein domains that transactivate or transrepress the RNA polymerase II. The regulatory system also has the ability to bind a small molecule (*e.g.*, a coinducer or a corepressor). The binding of the small molecule to the regulatory protein molecule (*e.g.*, a transcription factor) results in either increased or decreased affinity for the sequence element. Both inducible and repressible systems can be developed using any of the inducer/transcription factor combinations by positioning the binding site appropriately with respect to the promoter sequence. Examples of previously described inducible/repressible systems include lacI, ara, Steroid-RU486, and ecdysone - Rheogene, *Lac (Cronin et al. Genes & Development 15: 1506-1517, 2001), ara (Khlebnikov et al..*, J Bacteriol. 2000 Dec;182(24):7029-34), ecdysone (Rheogene, www.rheogene.com), RU48 (steroid, Wang XJ, Liefer KM, Tsai S, O'Malley BW, Roop DR., Proc Natl Acad Sci U S A. 1999 Jul 20;96(15):8483-8), *tet* promoter (Rendal *et al.*, Hum Gene Ther. 2002 Jan;13(2): 335-42. and Larnartina *et al.*, Hum Gene Ther. 2002 Jan; 13(2):199-210), or a promoter disclosed in WO 00/63364, filed April 19, 2000.

Nuclear transcription vectors for the generation of a nuclear dsRNA expression library

[0111] Nuclear transcription vectors for use in library based screening approaches to identify nucleic acids that modulate cell function, gene expression, or the biological activity of a target polypeptide are designed such that the target sequence is flanked on one end by an RNA polymerase II promoter (for example, the HCMV-IE promoter) and on the other end by a different RNA polymerase II promoter (for example, the SCMV promoter). Other promoters that can be used include other RNA polymerase II promoters, an RNA polymerase I promoter, an RNA polymerase III promoter, a mitochondrial RNA polymerase promoter, or a T7 or SP6 promoter in the presence of T7 or SP6 RNA polymerase, respectively, containing a nuclear localization signal. Bacteriophage or viral promoters may also be used. The promoters are regulated transcriptionally (for example, using a tet ON/OFF system (Forster *et al.*, *supra;* Liu *et al.*, *supra*; and Gatz, *supra*) such that they are only active in either the presence of a transcription-inducing agent or upon the removal of a repressor. A single chromosomal integrant is selected for, and transcription is induced in the cell to produce the nuclear dsRNA.

[0112] Those vectors containing a promoter recognized by RNA Pol I, RNA Pol II, or a viral promoter in conjunction with co-expressed proteins that recognize the viral promoter, may also contain optional sequences located between each promoter and the inserted cDNA. These sequences are transcribed and are designed to prevent the possible translation of a transcribed cDNA. For example, the transcribed RNA is synthesized to contain a stable stem-loop structure at the 5' end to impede ribosome scanning. Alternatively, the exact sequence is irrelevant as long as the length of the sequence is sufficient to be detrimental to translation initiation (*e.g.*, the sequence is 200 nucleotides or longer). The RNA sequences can optionally have sequences that allow polyA addition, intronic sequences, an HIV REV binding sequence, Mason-Pfizer monkey virus constitutive transport element(CTE) (U.S. 5,880,276, filed April 25, 1996), and/or self splicing intronic sequences.

[0113] To generate dsRNA, two promoters can be placed on either side of the target sequence, such that the direction of transcription from each promoter is opposing each other. Alternatively, two plasmids can be cotransfected. One of the plasmids is designed to transcribe one strand of the target sequence while the other is designed to transcribe the other strand. Single promoter constructs may be developed such that two units of the target sequence are transcribed in tandem, such that the second unit is in the reverse orientation with respect to the other. Alternate strategies include the use of filler sequences between the tandem target sequences.

Cytoplasmic transcription vectors for the generation of a cytoplasmic dsRNA expression library

[0114] Cytoplasmic transcription vectors for use in library based screening approaches to identifying nucleic acids that modulate cell function, gene expression, or the biological activity of a target polypeptide in a cell using PTGS are made according to the following method. This approach involves the transcription of a single stranded RNA template (derived from a library) in the nucleus, which is then transported into the cytoplasm where it serves as a template for the transcription of dsRNA molecules. The DNA encoding the ssRNA is integrated at a single site in the target cell line as described for the nuclear RNA expression library, thereby ensuring the synthesis of only one species of dsRNA in a cell, each cell expressing a different dsRNA species.

[0115] A desirable approach is to use endogenous polymerases such as the mitochondrial polymerase in animal cells or mitochondrial and chloroplast polymerases in plant cells for cytoplasmic and mitochondrial (e. g., chloroplast) expression to make dsRNA in the cytoplasm. These vectors are formed by designing expression constructs that contain mitochondrial or chloroplast promoters upstream of the target sequence. As described above for nuclear transcription vectors, dsRNA can be generated using two such promoters placed on either side of the target sequence, such that the direction of transcription from each promoter is opposing each other. Alternatively, two plasmids can be cotransfected. One of the plasmids is designed to transcribe one strand of the target sequence while the other is designed to transcribe the other strand. Single promoter constructs may be developed such that two units of the target sequence are transcribed in tandem, such that the second unit is in the reverse orientation with respect to the other. Alternate strategies include the use of filler sequences between the tandem target sequences.

[0116] Alternatively, cytoplasmic expression of dsRNA for use in library based screening approaches is achieved by a single subgenomic promoter opposite in orientation with respect to the nuclear promoter. The nuclear promoter generates one RNA strand that is transported into the cytoplasm, and the singular subgenomic promoter at the 3' end of the transcript is sufficient to generate its antisense copy by an RNA dependent RNA polymerase to result in a cytoplasmic dsRNA species.

Target cell line development for use with cytoplasmic dsRNA expression libraries

[0117] The target cell line, using the vector containing the G418 cassette, GFP, and *loxP* site is designed as described above.

Development of a cytoplasmic dsRNA expression library

[0118] Double stranded RNA expression libraries are generated by inserting cDNA or randomized sequences (as described herein) into an expression vector containing a single nuclear promoter (RNA polymerase I, RNA polymerase II, or RNA polymerase III), which allows transcription of the insert sequence. It is desirable that this nuclear promoter activity is regulated transcriptionally (for example, using a tet ON/OFF system described, for example, by Forster *et al.*, *supra*; Liu *et al.*, *supra*; and Gatz, *supra*), such that the promoters are only active in either the presence of a transcription-inducing agent or upon the removal of a repressor. This ensures that transcription is not induced until episomal copies of the vector(s) are diluted out. Vectors also contain a selectable marker, such as the hygromycin resistance gene, and a *loxP* site. The expression vectors are integrated into the target cell line by methods previously described in this application using Cre recombinase (other site-specific recombinative strategies can be employed, as described previously).

[0119] At two days post-transfection, cells are subjected to hygromycin selection using concentrations established in kill curve assays. Surviving cells are cultured in hygromycin to select for cells bearing integrated vectors and to dilute out episomal copies of the vector (s). At this point transcription is induced, and a single stranded RNA (ssRNA) species derived from the insert sequence is transcribed in the nucleus from the nuclear promoter in the inserted vector. The insert is designed such that the insert sequences in the transcript are flanked by bidirectional promoters of RNA bacteriophages (for example, Qbeta or MS2, RNA dependent RNA polymerase promoters) or cytoplasmic viral RNA-dependent RNA polymerase promoter sequences (for example, those of Sindbis or VEEV subgenomic promoters). The nuclear transcript is translocated to the cytoplasm where it acts as a template for dsRNA by an RNA dependent RNA polymerase, which may be provided through co-transfection of a vector that encodes an RNA-dependent RNA polymerase. Alternatively, an integrated copy of the polymerase may be used.

Example 2: Generation of templates for *in vitro* transcription of dsRNA for non-library based approaches for identification of nucleic acids using PTGS

[0120] Nucleic acid sequences that modulate cell function, gene expression in a cell, or the biological activity of a target polypeptide in a cell may also be identified using non-library based approaches involving PTGS. A single known nucleic acid sequence encoding a polypeptide with unknown function or a single nucleic acid fragment of unknown sequence and/or function can be made into a double stranded RNA molecule. This dsRNA is then transfected into a desired cell type and assayed for modulations in cell function, gene expression in the cell, or the biological activity of a target polypeptide in the cell, using methods described herein. A modulation in cell function, gene expression in the cell, or the biological activity of a target polypeptide in the cell identifies the nucleic acid of the dsRNA as a nucleic acid the modulates the specific cell function, gene expression, or the biological activity of a target polypeptide. This non-library based approach to nucleic acid identification is desirably used under conditions that inhibit an interferon response or dsRNA stress response. Such conditions are described in detail below.

[0121] Nucleic acid fragments generated, for example, by PCR or restriction endonuclease digestion, encoding the respective target sequences were used as templates for *in vitro* transcription reactions. PCR fragments are superior to plasmid templates for the synthesis of discrete sized RNA molecules. The PCR fragments encoded at least 20-50 or 100 to 1000, for exam-

ple, 500 to 600 nucleotides (nts) of the target sequence and were derived from the target mRNA. Known target sequences were obtained from GenBank and or other DNA sequence databases. Target sequences were also obtained from cellular RNAs that were generated into cDNAs to create a number of different dsRNA molecules. Accordingly, it is possible that the sequence and/or function of the target sequence was not known at the time the dsRNA was generated.

[0122] Templates for sense target RNAs were generated by placing the bacteriophage T7 promoter at the 5' end of the target coding strand while antisense RNA templates contained the T7 promoter at the 5' end of the non-coding strand. This was achieved by encoding the T7 promoter at the 5' ends of the respective PCR primers. Alternatively SP6 promoters, or a combination of SP6 and T7 promoters may be used.

[0123] PCR was performed by conventional methods. The use of both PCR templates in equimolar amounts in an *in vitro* transcription reaction resulted in primarily dsRNA. The use of two separate fragments has been found to be superior to the use of one PCR fragment containing two T7 promoters, one located at each end of the target sequence, presumably due to transcription interference that occurs during transcription of the dual promoter template. Following PCR amplification, the DNA was subjected to Proteinase K digestion and phenol-chloroform extraction to remove contaminating RNases. Following ethanol precipitation, the DNA was resuspended in RNase-free water at a concentration of 1 to 3 μg/μl.

[0124] As an alternative to phenol-chloroform extraction, DNA can be purified in the absence of phenol using standard methods such as those described by Li *et al.* (WO 00/44914, filed January 28, 2000). Alternatively, DNA that is extracted with phenol and/or chloroform can be purified to reduce or eliminate the amount of phenol and/or chloroform. For example, standard column chromatography can be used to purify the DNA (WO 00/44914, filed January 28, 2000).

Example 3: *In vitro* RNA transcription and RNA analysis

[0125] *In vitro* transcription reactions are carried out using the Riboprobe Kit (Promega Corp.), according to the manufacturer's directions. The template DNA is as described above. Following synthesis, the RNA is treated with RQ1 DNase (Promega Corp.) to remove template DNA. The RNA is then treated with Proteinase K and extracted with phenol-chloroform to remove contaminating RNases. The RNA is ethanol precipitated, washed with 70% ethanol, and resuspended in RNase-free water. Aliquots of RNA are removed for analysis and the RNA solution is flash frozen by incubating in an ethanol-dry ice bath. The RNA is stored at -80°C.

[0126] As an alternative to phenol-chloroform extraction, RNA can be purified in the absence of phenol using standard methods such as those described by Li *et al.* (WO 00/44914, filed January 28, 2000). Alternatively, RNA that is extracted with phenol and/or chloroform can be purified to reduce or eliminate the amount of phenol and/or chloroform. For example, standard column chromatography can be used to purify the RNA (WO 00/44914, filed January 28, 2000).

[0127] dsRNA is made by combining equimolar amounts of PCR fragments encoding antisense RNA and sense RNA, as described above, in the transcription reaction. Single stranded antisense or sense RNA is made by using a single species of PCR fragment in the reaction. The RNA concentration is determined by spectrophotometric analysis, and RNA quality is assessed by denaturing gel electrophoresis and by digestion with RNase T1, which degrades single stranded RNA.

[0128] An mRNA library is produced using Qbeta bacteriophage, by ligating the mRNAs to the flank sequences that are required for Qbeta replicase function (Qbeta flank or Qbeta flank plus P1), using RNA ligase. The ligated RNAs are then transformed into bacteria that express Qbeta replicase and the coat protein. Single plaques are then inoculated into fresh bacteria. All plaques are expected to carry transgene sequences. Each plaque is grown in larger quantities in bacteria that produce the Qbeta polymerase, and RNA is isolated from the bacteriophage particles. Alternatively, if the Qbeta flank plus P1 is used to generate the library (*e.g.*, P1=MS2, VEEV, or Sindbis promoter sequences), these vectors can be used to carry out the *in vitro* transcription along with the cognate polymerase. The *in vitro* made dsRNA is then used to transfect cells.

RNA delivery

[0129] *In vitro* made dsRNA is directly added to the cell culture medium at concentrations ranging from 50 μg/ml to 500 μg/ml. Uptake of dsRNA is also facilitated by electroporation using those conditions required for DNA uptake by the desired cell type. RNA uptake is also mediated by lipofection using any of a variey of commercially available and proprietary cationic lipids, DE-AE-dextran-mediated transfection, microinjection, protoplast fusion, calcium phosphate precipitation, viral or retroviral delivery, or biolistic transformation. The RNA is naked RNA or a local anesthetic RNA complex. Modulation of cell function, gene expression, or polypeptide biological activity is then assessed in the transfected cells.

[0130] Some dsRNA sequences, possibly in certain cell types and through certain delivery methods, may result in an interferon response. During the screening methods of the present invention, induction of an interferon response is not desired, as this would lead to cell death and possibly to the prevention of gene silencing.

[0131] One of the components of an interferon response is the induction of the interferon-induced protein kinase PKR. Suppression of the interferon response and/or the PKR response is desired in the target cells.

The dsRNA delivery methods described herein are performed such that an interferon response or dsRNA stress response is not included. It is recognized, however, that certain conditions might present with an induction of the interferon response. To prevent such a response, a number of other strategies may be employed with any of the above described screening methods to identify a nucleic acid that modulates cell function, gene expression, or the polypeptide biological activity of a cell, as described herein.

[0132] To prevent an interferon response, interferon and PKR responses are silenced in the target cells using a dsRNA species directed against the mRNAs that encode proteins involved in the response. Desirably interferon response promoters are silenced using dsRNA. Alternatively, the expression of proteins that bind the interferon response element is abolished using dsRNA techniques.

[0133] In an alternative strategy, interferon and PKR knockout cell lines are created through approaches utilizing expression cassettes that encode an antisense RNA and ribozymes directed to the cellular mRNAs that encode the proteins involved in the response. Knockout cells are created by standard gene knockout technologies using homologous recombination to alter target sequences, using homologous DNA alone, or as complexes of RecA protein and single stranded DNA homologous to the target sequence(s). Interferon response element (IRE) sequences, sequences that encode transcription factors that bind IRE sequences, the promoter and/or gene sequences that encode proteins in the PKR and interferon response pathways are molecules that are targeted for knockout.

[0134] If desired, proteins involved in gene silencing such as Dicer or Argonaut can be overexpressed or activated to increase the amount of inhibition of gene expression (Beach *et al.*, WO 01/68836, filed March 16, 2001).

Example 4: Cytoplasmic transcription vectors for non-library based approaches to nucleic acid identification using PTGS

[0135] Double stranded RNA molecules for use in non-library based methods for the identification of nucleic acids that modulate cell function, gene expression of a target nucleic acid, or target polypeptide biological activity in a cell can also be generated through the use of cytoplasmic transcription vectors. Such vectors are generated as now described.

[0136] The PCR fragments generated for *in vitro* transcription templates, as described above, are inserted into a cloning vector containing one T7 promoter located just outside the polylinker region. Such a vector is pZERO blunt (Promega Corp.). The PCR fragment is cloned into a restriction site in the polylinker in such a way that the fragment's T7 promoter is distal to the vector's promoter. The resulting vector contains the target sequence flanked by two T7 promoters; transcription from this vector occurs in converging directions. Convergent transcription is desired for these intracellular vectors, due to the uncertainty of getting sense and antisense vectors into the same cell in high enough and roughly equivalent amounts. In addition, the local concentration of antisense and sense RNAs with respect to each other is high enough to enable dsRNA formation when the dual promoter construct is used.

[0137] A hygromycin resistance cassette is cloned into the pZERO blunt vector as well. The hygromycin resistance cassette contains the hygromycin resistance gene under the control of the Herpes Simplex Virus (HSV) thymidine kinase promoter and the SV40 polyadenlyation signal. The cassette is in a plasmid vector and is flanked at both ends by a polylinker region enabling ease of removal and subsequent cloning. Hygromycin selection was chosen because of the rapidity of death induced by hygromycin as well as extensive in-house experience with hygromycin selection. Alternatively, other selection methods known to those skilled in the art may be used.

[0138] The vectors are transfected into the desired cells using standard transformation or transfection techniques described herein, and the cells are assayed for the ability of the dsRNA molecules encoded by the vectors to modulate cell function, gene expression of a target nuclec acid, or the biological activity of a target polypeptide, as described herein.

Example 5: Analysis of RNA from transfected cells

[0139] Regardless of whether a library based screening approach or a non-library based approach was used to identify nucleic acid sequences, in order to measure the level of dsRNA effector molecule within the cell, as well as the amount of target mRNA within the cell, a two-step reverse transcription PCR reaction was performed with the ABI PRISM™ 7700 Sequence Detection System. Total RNA was extracted from cells transfected with dsRNA or a plasmid from a dsRNA expression library using Trizol and DNase. Two to three different cDNA synthesis reactions were performed per sample; one for human GAPDH (a housekeeping gene that should be unaffected by the effector dsRNA), one for the target mRNA, and/or one for the sense strand of the expected dsRNA molecule (effector molecule). Prior to cDNA synthesis of dsRNA sense strands, the RNA sample was treated with T1 RNase. The cDNA reactions were performed in separate tubes using 200 ng of total RNA and primers specific for the relevant RNAs. The cDNA products of these reactions were used as templates for subsequent PCR reactions to amplify GAPDH, the target cDNA, and/or the sense strand copied from the dsRNA. All RNA was quantified relative to the internal control, GAPDH.

Example 6: Target sequence identification

[0140] To identify the target sequence affected by a dsRNA, using any of the above-described methods, DNA is extracted from expanded cell lines (or from the transfected cells if using a non-integrating dsRNA system) according to methods well known to the skilled artisan. The dsRNA encoding sequence of each integrant (or non-integrated dsRNA molecule if using a non-library based method) is amplified by PCR using primers containing the sequence mapping to the top strand of the T7 promoter (or any other promoter used to express the dsRNA). Amplified DNA is then cloned into a cloning vector, such as pZERO blunt (Promega Corp.), and then sequenced. Sequences are compared to sequences in GenBank and/or other DNA databases to look for sequence identity or homology using standard computer programs. If the target mRNA remains unknown, the mRNA is cloned from the target cell line using primers derived from the cloned dsRNA by established techniques (Sambrook *et al.*, *supra*). Target validation is then carried out as described herein.

[0141] In the stably integrated dsRNA expression system described above, despite efforts to reduce negative position effects, inefficient dsRNA synthesis by PCR methods may occur. This can be circumvented by rescuing the integrated cDNA or randomized nucleic sequences into replicating plasmids. Rescued plasmids are amenable to amplification in bacteria and to sequencing. Rescue is achieved by re-transfecting the population of cells transfected with the dsRNA expression library with the rescue plasmid and a plasmid encoding Cre recombinase. The rescue plasmid carries a bacterial origin of replication, a bacterial antibiotic selection marker, an SV40 origin of replication, and an SV40 T antigen expression cassette, as well as *loxP* sites positioned as an inverted repeat to allow Cre-mediated double recombination. The SV40-based origin of replication in the rescue plasmid allows amplification of rescued sequences in the integrated cells. Following rescue, higher levels of transcription are anticipated, thereby favoring dsRNA formation. The cells are then screened for modulations in cell function, target nucleic acid expression, or target polypeptide biological activity changes as described herein.

Example 7: Prevention of an interferon response during gene silencing

[0142] As discussed above, during the above-described screening methods, induction of an interferon response is not desired, as this would lead to cell death, anti-proliferative responses, and possibly to prevention of gene silencing. One of the components of an interferon response is the induction of the interferon-induced protein kinase PKR. Suppression of the interferon response and/or the PKR response is desired in the target cells. The dsRNA delivery methods described herein are performed such that an interferon response is not included. It is recognized, however, that certain conditions might present with an induction of the interferon response. To prevent such a response, a number of other strategies may be employed with any of the above described screening methods to identify a nucleic acid that modulates cell function, gene expression, or the polypeptide biological activity of a cell, as described herein.

[0143] To prevent an interferon response in a system involving stable integration of the nucleic acid containing the dsRNA expression cassette, the vectors used to generate either the *loxP* integrant or the vector that encodes the dsRNA expression cassette are designed to contain sequences that encode proteins that block the PKR response, such as the Vaccinia virus protein E3 (Romano *et al.*, Molecular and Cellular Biology 18: 7304-7316, 1998; Accession No. M36339), or a cellular protein $P58^{IPK}$, which the influenza virus mobilizes to block PKR (Gale *et al.*, Microbiology and Molecular Biology Reviews 64:239-280, 2000; Accession No. XM_032882). Several other viral proteins have also been identified (*e.g.*, Hepatitis C E2; Accession No. S72725) and may be similarly used. These proteins can be expressed in the desired cell types or in animals through the use of any of a number of commercially available mammalian expression vectors or vertebrate expression vectors. Such vectors can be obtained from a number of different manufacturers including Invitrogen (Carlsbad, CA) Promega ((Madison, WI), or Clontech (Palo Alto, CA). An example of such a vector is the pCI-neo Mammalian Expression Vector from Promega.

[0144] Regardless of whether nucleic acid encoding a dsRNA is stably integrated into a chromosome or is not integrated into a chromosome, the following methods may be used to prevent an interference response in any of the screening methods of the present invention. In one example of an interferon avoidance strategy, interferon and PKR responses are silenced in the target cells using a dsRNA species directed against the mRNAs that encode proteins involved in the response. Desirably interferon response promoters are silenced using dsRNA. Alternatively, the expression of proteins that bind the interferon response element is abolished using dsRNA techniques.

[0145] In an alternative strategy, interferon and PKR knockout cell lines are created through approaches utilizing expression cassettes that encode an antisense RNA and ribozymes directed to the cellular mRNAs that encode the proteins involved in the response. Knockout cells are created by standard gene knockout technologies using homologous recombination to alter target sequences, using homologous DNA alone, or as complexes of RecA protein and single stranded DNA homologous to the target sequence(s). Interferon response element (IRE) sequences, sequences that encode transcription factors that bind IRE sequences, the promoter and/or gene sequences that encode proteins in the PKR

and interferon response pathways are molecules that are targeted for knockout.

[0146] In yet another alternative, chimeric oligonucleotides may be used to alter target sequences. Methods for inhibiting expression of polypeptides through chimeric oligonucleotides are well known in the art (Igoucheva and Yoon, Human Gene Therapy 11:2307-2312, 2000).

Example 8: Functional screening for cell invasion

[0147] Cell invasion is one cell function that may be evaluated in the search for novel genes that are modulated using the methods described herein. Matrigel, a biological extracellular matrix, has properties similar to that of a reconstituted basement membrane and has been used to measure the invasive potential of tumor cells (Platet and Garcia, *supra*). Cells transfected with randomized or cDNA libraries that have been cloned into PTGS vectors are monitored for their capacity to invade matrigel invasion chambers. Cells that have taken up sequences unrelated to invasion invade the matrigel as efficiently as vector-transfected control cells. Cells experiencing PTGS of genes that are involved in cell invasion invade much less efficiently. If the dsRNA expression cassette is stably integrated in a chromosome, these cells are retrieved and second and third rounds of selection are carried out in order to isolate specific nucleic acid sequences relevant to cell invasion. The effect of these sequences on invasion is ultimately confirmed by their ability to block the formation of tumors in animal models.

[0148] Several human cell lines, for example, MDA-MB-231, used by Platet and Garcia (*supra*), SKBr3, and MCF-7ADR, a more metastatic variant of MCF-7. MDA-MB-231 breast cancer cells (obtained from the American Type Culture Collection) are also transfected with cDNA libraries or randomized nucleic acid libraries constructed into the vectors described above. Desirably all cells in this assay contain a single copy of a transfected gene, as described above.

[0149] Cells cultured in commercially available 24- or 96-well formatted systems are used to carry out the cell invasion assay. As this screening protocol relies upon repeated rounds of selection, it may be desirable to keep the cell numbers in each well low enough that enrichment is seen in each succeeding round, yet high enough to recover sufficient cells to culture within a reasonable time period. Therefore, culture conditions that result in invasion by greater than 50% of the cells and that still permit recovery from the surface of the matrigel are made optimal. Non-invasive (NIH3T3 cells) or poorly invasive (MCF7) cell lines are analyzed in parallel as negative controls for invasion.

[0150] Initially, triplicate cultures of half-log order dilutions from 10^2 to 10^6 cells per well are plated. Cells are then recovered by "scrubbing" with a sterile cotton swab in fresh culture media and are seeded into 96-well plates. The number of invasive cells in the matrigel is quantified using either an MTT-based assay (Sasaki and Passaniti, Biotechniques 24:1038-1043, 1998) or a fluorescent indicator (Gohla *et al.*, Clin. Exp. Metastasis 14:451-458, 1996).

[0151] Once the appropriate cell densities for the assay have been empirically determined, stable transfected cells are plated in the matrigel cell invasion chambers. Each experiment includes the following controls: a sample of untransfected cells as a reference culture; untransfected cells treated with anti-invasive chemotherapeutic agents, such as taxol or doxorubicin, as a positive control for inhibition of invasion; cells transfected with empty vectors to confirm that the vector alone had no effects on invasion; and cells transfected cells with genes that are known to block invasion in this assay, such as estrogen receptor-α or TIMP-2 (Kohn *et al.*, Cancer Research 55:1856-1862, 1995; and Woodhouse *et al.*, Cancer (Supplement) 80:1529-1536, 1997).

[0152] Cells that fail to invade the matrigel are removed from each well to the corresponding wells of a 96-well plate and cultured until macroscopic colonies are visible. It is important to collect cells at more than one time point after plating, since the time it takes for PTGS to be effective may vary, and it may be that different genes are active at different times after plating. Once the cells are transferred to 96-well plates, they are diluted out and taken through successive rounds of re-screening in the invasion assay in order to expand and isolate cell lines with altered invasive ability. As the population becomes more and more enriched for cells with a non-invasive phenotype, the reduction in invasive cells in the matrigel can be better quantified via MTT or fluorescence assays. Ultimately, a large panel of cloned double-stable cell lines is generated.

[0153] This assay can also be carried out with cells into which a dsRNA is not stably integrated into a chromosome. The assay is conducted essentially as described above except that multiple rounds of selection and re-screening are not necessary since the cell is transfected with only one dsRNA species. Thus, the target(s) of the PTGS event is readily identifiable using the cloning and sequencing techniques described above.

Example 9: Downregulation of HIV using HIV-derived dsRNA and inhibitors of the interferon response pathway

[0154] During the course of HIV infection, the viral genome is reverse transcribed into a DNA template that is integrated into the host chromosome of infected dividing cells. The integrated copy is now a blueprint from which more HIV particles are made. Several cell lines that contain integrated copies of a defective HIV genome, HIVgpt (strain HXB2) have been created. The HIVgpt genome contains a deletion of the HIV envelope gene; all other HIV proteins are encoded. The plasmid used to create these cell lines, HIVgpt, was obtained from the

AIDS Research and Reference Reagent Program Catalog. Stably integrated cell lines were made with human rhabdomyosarcoma (RD) cells. The lines were made by transfecting cells with the plasmid followed by selection of cells in mycophenolic acid. The HIVgpt genome encodes a mycophenolic acid (MPA) resistance gene in place of the envelope gene and thereby confers resistance to MPA. Cells resistant to MPA were clonally amplified. The media from the cultured clonally expanded cells was assayed for the presence of p24 (an HIV gag polypeptide that is secreted extracellularly). All cell lines were positive for p24, as assessed using a p24 ELISA assay kit (Coulter, Fullerton, CA). The cell lines also make non-infectious particles which can be rescued into infectious particles by co-expression of an HIV envelope protein.

[0155] The HIVgpt cell lines are used as a model system with which to downregulate HIV expression via PTGS. Plasmids encoding a 600 nt sense RNA, a 600 nt antisense RNA, or a 600 bp double stranded RNA (dsRNA), mapping to the same coordinates of the gag gene of HIV strain HXB2 are used to transfect cells (the map from which the coordinates are based is found at GenBank Accession number K03455, HIV(HXB2), complete genome, and the gag RNAs used in this study map to coordinates 901-1500). Expression of the RNAs is from T7 RNA polymerase promoter(s) located at the 5' end of the gag sense strand, at the 5' end of the antisense strand, or at converging positions at the 5' ends of both the sense and anti-sense strands, respectively. These encoded RNAs are not designed to be able to make protein (*i.e.*, they do not have a cap, a poly A tail, or the native initiation codon). Transcription of the RNAs is catalyzed by T7 RNA polymerase, provided from a second co-transfected T7 RNA polymerase expression plasmid. Control plasmids expressing a similar sized sense RNA, antisense RNA, and dsRNA derived from the gD gene of an HSV2 genome are included as experimental controls (the map from which the coordinates are based is found at GenBank Accession number K01408, HSVgD2 gene, and the HSVgD RNAs used in this study map to coordinates 313-872).

[0156] Cells used in these studies are transfected with an expression plasmid encoding a gene product known to interfere with the dsRNA induced interferon response or with the PKR response, as described above. The cells are transfected with lipofectamine (Gibco-BRL) as a transfecting agent according to the manufacturer's instructions.

[0157] Two days after transfection, the cells are harvested and seeded into six-well plates and cultured to approximately 80 to 90% confluence. Cells are co-transfected with the T7 RNA polymerase expression plasmid and one of the RNA expression plasmids, such that one well of cells receives the T7 RNA polymerase expression plasmid and the gag sense RNA expression plasmid; one well of cells receives the T7 RNA polymerase expression plasmid and the gag antisense RNA expression plasmid; one well of cells receives the T7 RNA polymerase expression plasmid and the gag dsRNA expression plasmid; one well of cells receives the T7 RNA polymerase expression plasmid and the HSVgd sense RNA expression plasmid; one well of cells receives the T7 RNA polymerase expression plasmid and the HSVgd antisense RNA expression plasmid; and one well of cells receives the T7 RNA polymerase expression plasmid and the HSVgD dsRNA expression plasmid. Transfection is again mediated through lipofectamine (Gibco-BRL). There also is a control group of cells receiving no RNA. The cells are monitored for p24 synthesis over the course of several weeks. The cells are assayed both by measuring p24 in the media of cells (using the p24 ELISA kit from Coulter, according to the manufacturer's instructions) and by immunostaining fixed cells for p24 using a rabbit polyclonal anti-p24 sera and anti-rabbit IgG that is FITC conjugated (Sigma). None of the gD RNAs specifically shut down p24 synthesis. The double stranded gag RNA significantly down regulates p24. The sense and antisense have only a modest effect on p24 synthesis and some of the effect is predicted to be through the ability of the sense and antisense gag RNAs to generate low levels of dsRNA species.

Example 10: Downregulation of PSA expression in human Rhabdomyosarcoma cells using intracellular expression of dsRNA

[0158] RD cells transiently expressing prostate specific antigen (PSA) were transfected with a T7 RNA polymerase expression vector and T7 RNA expression vectors expressing PSA dsRNA, PSA sense RNA, PSA antisense RNA, or control RNAs. The ability of the expressed RNAs to downregulate PSA expression was assessed, as described further below.

Creation of a transient PSA expression line

[0159] The ability to downregulate expression of PSA following the expression of a PSA specific double-stranded RNA (dsRNA) was demonstrated in a human rhabdomyosarcoma cell line. Since available PSA cell lines are difficult to work with (*i.e.*, they are hard to transfect, and the cells tend to clump), a human cell line transiently expressing PSA was created. To create these cells, human rhabdomyosarcoma cells were transiently transfected with a PSA plasmid-based expression vector, under conditions that result in transfection of greater than 95% of the cells. Transfection was mediated with lipofectamine transfecting reagent (Gibco-BRL) according to the manufacturer's instructions. Expression of PSA was directed by the HCMV-IE promoter and the SV40 polyadenylation signal (Fig. 2). PSA expression was measured in the supernatant of transfected cells using a PSA ELISA kit (Oncogene Science Diagnostics, Cambridge, MA). No PSA was detected in the untransfected parental cells while PSA was abundantly ex-

pressed in cells receiving the PSA expression vector.

Downregulation of PSA expression

[0160] The PSA expressing cell line was used as a model system with which to demonstrate the downregulation of PSA protein levels by PTGS. In these studies, plasmids encoding an approximately 600 nt sense RNA, a 600 nt antisense RNA or a 600 nt dsRNA derived from a PSA cDNA were used to transfect the PSA expressing cell line (Fig. 2). Expression of the RNAs was from a T7 RNA polymerase promoter(s) located at the 5' end of the PSA sense strand, at the 5' end of the PSA antisense strand, or at converging positions at the 5' ends of both the sense and antisense strands respectively (Fig. 2). These encoded RNAs are not designed to be able to make protein (they do not have a cap, or a poly A tail). Transcription of the RNAs was catalyzed by T7 RNA polymerase, provided by a co-transfected T7 RNA expression plasmid. Control plasmids expressing similar sized sense RNA, antisense RNA, and dsRNA derived from the glycoprotein D (gD) gene of an Herpes simplex 2 (HSV-2) genome, as described above were included as experimental controls.

[0161] Cells used in these studies can optionally be transfected with an expression plasmid encoding a gene product known to interfere with the dsRNA induced interferon response or with the PKR response, as described above. The cells are transfected with lipofectamine (Gibco-BRL) as a transfecting agent according to the manufacturer's instructions.

[0162] Human rhabdomyosarcoma cells were seeded into six-well plates and cultured to approximately 80 to 90% confluence. The cells were co-transfected with (A) the PSA expression plasmid; (B) one of the T7 RNA expression plasmids; and (C) the T7 RNA polymerase expression plasmid, such that all PSA expressing cells were transfected with the T7 RNA polymerase expression plasmid and one of the following: the T7 sense PSA RNA expression construct, the T7 antisense PSA RNA expression construct, the T7 dsRNA PSA expression construct, the sense HSVgD RNA expression construct, the antisense HSVgD expression construct, or the dsRNA HSVgd expression construct. Cells received identical amounts of the PSA expression plasmid and the T7 RNA expression plasmid. The amounts of the T7 RNA expression plasmids were also constant amongst the transfected cells. Total DNA per transfection was held constant at 2.5 μg DNA per one well of a six-well plate. In those transfections where there was no T7 RNA expression plasmid, an inert plasmid DNA was used as filler DNA. Transfection was mediated by lipofectamine (Gibco-BRL) according to the manufacturer's instructions. There was also a control group of untransfected cells, as well as an untreated PSA control group of cells transfected with only the PSA expression plasmid in combination with the T7 RNA polymerase expression plasmid.

[0163] PSA-expressing cells that were not transfected with the T7 RNA expression plasmid, as well as cells transfected with the T7 HSV2-gDRNA expression plasmid all expressed PSA abundantly and at comparable levels. Cells transfected with the sense, antisense, and ds PSA RNA expression plasmids all exhibited varying degrees of inhibition of PSA expression. A 5%-10% reduction in expression was seen in cells expressing the PSA sense RNA, a 50% reduction was seen in cells expressing the PSA antisense RNA and greater than 95% reduction was seen in the cells expressing the PSA dsRNA (Fig. 3). The inhibition was seen within two days after transfection and continued up until the last time point taken (one month later) at which point PSA levels were beginning to decline in the untreated cells and the experiment was terminated. The untreated PSA controls as well as cells transfected with the T7 HSV2-gD control RNA expression plasmids all expressed PSA abundantly and at comparable levels (Fig. 3), indicating the specificity of dsRNA effectors to silence gene expression. During the one of month culture, cells were expanded into larger cultures at routine intervals.

[0164] Although the PSA specific dsRNA induced significant inhibition of PSA expression, antisense and sense PSA RNAs also induced some level of inhibition. Antisense PSA RNA has the potential to form dsRNA by annealing with PSA mRNA. Therefore the inhibition seen with antisense RNA may be explained by both an antisense mechanism and a dsRNA induced inhibition. A critical intracellular concentration of both antisense RNA and mRNA is required to generate dsRNA. Since much less dsRNA is made in the antisense RNA expressing cells relative to those cells designed to make dsRNA, a lesser inhibition of PSA in the antisense RNA expressing cells is expected if the threshold dsRNA levels required for efficient silencing have not been reached in those cells. We have also demonstrated that a small amount of antisense RNA can be found in cells transfected with our expression vectors (approximately 0.2% the amount of mRNA steady state levels). Antisense expression is presumably driven by cryptic promoters on the non-coding plasmid DNA strand. The observed sense RNA inhibition could therefore also involve a dsRNA molecule. RNA from transfected but untreated cells could also be analyzed to determine if the low level expression of antisense RNAs in these cells results in the production of detectable dsRNA species. Some low level expression of PSA occurred in cells expressing PSA dsRNA. It is likely that some percentage of cells did not take up the dsRNA expression cassette or that the threshold levels of dsRNA were not reached in some cells. No cellular toxicity was seen with any of the dsRNAs generated by the RNA expression vectors suggesting that cytoplasmic expression of dsRNA does not induce the interferon response. In contrast, cell death is induced when certain concentrations of *in vitro* produced dsRNA is delivered to cells via transfection with certain cationic lipids.

[0165] In summary, these results indicate that (i) PSA derived dsRNA is much more efficient than PSA antisense RNA in down-regulating PSA expression, (ii) the down-regulation is sequence specific; only the PSA derived dsRNA and not the control HSV-2 derived dsRNA induced down-regulation of PSA, and (iii) there is no toxicity associated with the cytoplasmic expression of long (600 bp) dsRNA molecules. Additionally, these experiments are the first demonstration of dsRNA mediated down-regulation of gene expression in a human cell line.

Example 11: Intracellular expression of dsRNA does not induce the type 1 interferon response (RNA stress response).

[0166] Human rhabdomyosarcoma (RD) cells were transfected with various dsRNA expression vectors such that dsRNA was transcribed in the transfected cells as described in Example 10. Transcription of dsRNA occurred in the cells within 24 hours after transfection and continued for the duration of the thirty day experiment. Cells and the supernatants from transfected cells were analyzed during the course of the experiment for any evidence of RNA stress response induction. No evidence of RNA stress response induction by intracellular expressed dsRNA was observed. RD cells have been shown by us to be responsive to type 1 interferon, both alpha and beta, and thus RD cells are capable of mounting an RNA stress response. In addition, positive controls were included in these experiments. A positive control for these experiments is a method of delivering dsRNA which induces the RNA stress response. All positive controls induced the RNA stress response. These experiments are described further below.

Assays performed to identify RNA stress response induction

[0167] The following assays were performed to measure the induction of an RNA stress response: TUNEL assay to detect apoptotic cells, ELISA assays to detect the induction of alpha, beta and gamma interferon, ribosomal RNA fragmentation analysis to detect activation of 2'5'OAS, measurement of phosphorylated eIF2a as an indicator of PKR (protein kinase RNA inducible) activation, proliferation assays to detect changes in cellular proliferation, and microscopic analysis of cells to identify cellular cytopathic effects. Apoptosis, interferon induction, 2'5' OAS activation, PKR activation, anti-proliferative responses, and cytopathic effects are all indicators for the RNA stress response pathway.

Transfection of cells

[0168] Approximately $7x10^5$ RD cells were seeded into individual wells of six-well plates. Cells were transfected when they reached about 90% confluency. Cells were transfected with a T7 RNA polymerase expression construct and a T7 dsRNA expression construct. The T7 dsRNA expression constructs encode converging T7 promoters located on either side of a 600 bp sequence (Fig. 2). Controls included cells transfected with the T7 RNA expression construct alone so that no dsRNA is made in these cells. Total DNA per transfection was held constant at 2.5 µg DNA per one well of a six-well plate. When the T7 RNA polymerase and T7 dsRNA expression vectors were both used, 1.25 µg of each DNA was used per transfection. In those transfections where there was no T7 dsRNA expression construct, inert filler DNA was used to bring the total DNA to 2.5 µg per transfection. Transfection was mediated using Lipofectamine (InVitrogen) according to the manufacturer's directions. The positive control transfections included poly(I)(C) RNA and *in vitro* transcribed dsRNA of 600 bp that were both complexed with Lipofectamine and transfected into cells. The cells were transfected with *in vitro* transcribed ssRNA complexed with Lipofectamine. 2.5 µg of each RNA was used per transfection. Other controls included untreated cells. Cells were kept in culture for one month by expanding into larger flasks as the cell numbers increased.

ELISA assays

[0169] Supernatants were removed from the transfected and untreated cells at time points of 1, 2, 7, 17, and 48 hours and every several days for up to one month after the 48 hour time point. Collected supernatants were stored at -80°C until they were analyzed for the presence of alpha, beta, and gamma interferon using commercially available ELISA kits. The Interferon-alpha ELISA kit was obtained from ENDOGEN (Rockford, IL), the Interferon-Beta ELISA kit was obtained from RD1 (Flanders, NJ), and the Interferon-gamma ELISA kit was obtained from R&D Systems (Minneapolois,MN). ELISAs were all performed according to the manufacturer's directions. Alpha, beta, and gamma interferon were not detected at increased levels in cells expressing intracellular dsRNA compared to the corresponding levels in untreated cells. However, considerable levels of beta interferon were found in cells transfected with poly (I)(C) or with *in vitro* transcribed dsRNA and ssRNA. Alpha and beta interferon induction are associated with induction of the RNA stress response.

TUNEL assay

[0170] Cells were stained for the presence of apoptotic nuclei using a commercially available kit, TdT FragEL, DNA Fragmentation Detection Kit, In Situ Apoptosis Assay from Oncogene (Boston, MA). Cells were stained according the manufacturer's directions. Cells were stained at 2 hours, 7 hours, 17 hours, 2 days, 3 days, 4 days, and 5 days after transfection. There was no evidence of apoptosis induced by intracellular expressed dsRNA at any of the time points analyzed. How-

ever, the majority of cells transfected with poly (I)(C) or with the *in vitro* transcribed dsRNA were apoptotic by 17 hours after transfection. No evidence of apoptosis was observed in the untreated cells or in cells transfected with ssRNA. Apoptosis is an end result of the induction of the RNA stress response pathway.

2'5'OAS activation

[0171] The activation of 2'5'OAS was determined by performing ribosomal RNA fragmentation analysis. Briefly, following transfection, total RNA was extracted from cells using standard procedures. RNA was extracted at the following time points: 2 hours, 7 hours, 17 hours, 48 hours, 3 days, 4 days, and 5 days after transfection. 5-10 µg RNA was analyzed for each sample. RNA samples were first denatured in formaldehyde/formamide RNA sample buffer at 65°C for 10 minutes prior to being electrophoresed through 0.5X TBE agarose gels. Ribosomal RNA was visualized by staining with ethidium bromide followed by ultraviolet transillumination. Ribosomal RNA fragmentation was observed in cells transfected with poly (I)(C) and with the *in vitro* transcribed dsRNA. No fragmentation was observed in the untreated control cells, cells transfected with ssRNA, or in cells expressing intracellular dsRNA. These results indicate that 2'5'OAS was not activated by dsRNA when it was made intracellularly. 2'5'OAS activation is associated with induction /activation of the RNA stress response pathway.

PKR activation

[0172] The activation of PKR was determined by measuring the levels of eIF2alpha phosphorylation. Briefly, cells were lysed at various times after transfection (2 hours, 7 hours, 19 hours, 48 hours, 3 days, 4 days, and 5 days after transfection) and analyzed for levels of phosphorylated and non-phosphorylated eIF2 alpha. The protocol for lysing cells can be found in the following reference: Zhang F. *et al.*, J. Biol. Chem. 276 (27):24946-58, 2001. This analysis was performed as described for detecting PKR phosphorylation except that antibodies specific for phosphorylated and non-phosphorylated eIF2alpha were used. These antibodies are available from Cell Signaling Technology (Beverly, MA).

Cytopathic effect and antiproliferative responses

[0173] Cytopathic effect was assayed by analyzing cells microscopically using a light microscope. Cells were analyzed at daily intervals throughout the course of the experiment. Cytopathic effect is defined as any or all of the following: cells detaching from surface of well/ flask, cells rounding up, an increased number of vacuoles in transfected cells with respect to the control untreated cells, or differences in morphology of cells with respect to the untreated control cells. No cytopathic effect was seen in those cells expressing dsRNA intracellularly. Severe cytopathic effects were seen in cells transfected with Poly (I)(C) or with dsRNA made *in vitro*. Cytopathic effect is associated with the RNA stress response.

[0174] Antiproliferative responses were assayed by measuring the division rate of cells. The division rate is determined by counting cell numbers using standard procedures. Cells were counted every few days for the duration of the experiment. No antiproliferative responses were seen in cells expressing dsRNA intracellulary. Antiproliferative responses are associated with the RNA stress response.

Summary of results

[0175] The results of the above assays indicate that intracellular expression of dsRNA does not induce the RNA stress response. The cells that were used for these experiments were competent for RNA stress response induction as was demonstrated by the ability of cationic lipid complexed poly(I)(C) and *in vitro* transcribed RNA to induce/activate all tested components of this response. In addition, the cells were found to be responsive to exogenously added interferon. These results imply that the cells used for these experiments are not defective in their ability to mount an RNA stress response and therefore can be used as predictors for other cells, both in cell culture and *in vivo* in animal models. This method described here, which does not induce the interferon stress response, has been found to induce PTGS. This method therefore provides a method to induce PTGS without inducing an undesired RNA stress response.

[0176] Although these results were generated using a vector that utilizes a T7 transcription system and therefore expresses dsRNA in the cytoplasm, the vector system can be changed to other systems that express dsRNA intracellularly. Similar results are expected with these expression systems. These systems include, but are not limited to, systems that express dsRNA or hairpin RNAs in the nucleus, in the nucleus followed by transport of the RNAs to the cytoplasm, or in the cytoplasm using non-T7 RNA polymerase based expression systems.

Example 12: Optimization of the concentrations and relative ratios of *in vitro* or *in vivo* produced dsRNA and delivery agent

[0177] The optimal concentrations and ratios of dsRNA to a delivery agent such as a cationic lipid, cationic surfactant, or local anesthetic can be readily determined to achieve low toxicity and to efficiently induce gene silencing using *in vitro* or *in vivo* produced dsRNA.

Summary of factors effecting nucleic acid/cationic lipid interactions

[0178] Cationic lipid DNA interactions are electrostatic. Electrostatic interactions are highly influenced by the ionic components of the medium. The ability to form stable complexes is also dependent upon the intermolecular interactions between the lipid molecules. At low concentrations, certain inter-lipid interactions are preferred; at higher lipid concentrations, rapid condensates are formed due to higher order interactions. Although local interactions are similar in both of these instances (*e.g.*, phosphoryl groups in the DNA and the charged cationic head group), the long range and inter-lipid interactions are substantially different. Similarly, structurally diverse variants can be obtained simply by changing the charge ratio of the complex by mixing varying amounts of cationic lipid with fixed concentrations of the nucleic acid or *vice versa*. This variation in the structure of the complexes is evidenced by altered physical properties of the complexes (*e.g.*, differences in octanol partitioning, mobility on density gradients, charge density of the particle, particle size, and transfectability of cells in culture and *in vivo*) (Pachuk *et al.* DNA Vaccines - Challenges in Delivery, Current Opinion in Molecular Therapeutics, 2(2) 188-198, 2000 and Pachuk *et al.*, BBA, 1468, 20-30, (2000)). Furthermore, different lipids, local anesthetics, and surfactants differ in their interactions between themselves, and therefore novel complexes can be formed with differing biophysical properties by using different lipids singularly or in combination. For each cell type, the following titration can be carried out to determine the optimal ratio and concentrations that result in complexes that do not induce the stress response or interferon response. At several of these concentrations PTGS is predicted to be induced; however, PTGS is most readily observed under conditions that result in highly diminished cytotoxicity.

Complex formation

[0179] dsRNA is either produced by *in vitro* transcription using the T7 promoter and polymerase or another RNA polymerase, such as an *E. coli RNA polymerase. dsRNA* can also be produced in an organism or cell using endogenous polymerases.

[0180] Concentrations of dsRNA, such as PSA-specific dsRNA, are varied from 1 pg to 10 μg. In some instances, 150 ng of a plasmid that encodes a reporter of interest (PSA) to be silenced may be comixed at a concentration between 10 ng and 10 μg. The concentration of cationic lipid, cationic surfactant, local anesthetic, or any other transfection facilitating agent that interacts with the nucleic acid electrostatically are varied at each of the dsRNA concentrations to yield charge ratios of 0.1 to 1000 (positive/ negative) (*i.e.*, the ratio of positive charge from lipids or other delivery agents to negative charge from DNA or RNA). The complexes are prepared in water or in buffer (*e.g.*, phosphate, HEPES, citrate, Tris-HCl, Tris-glycine, malate, *etc.* at pH values that range from 4.0 to 8.5), may contain salt (*e.g.*, 1 - 250 mM), and may contain glycerol, sucrose, trehalose, xylose, or other sugars (*e.g.*, mono-, di-, or polysaccharide). The mixture is allowed to sit at room temperature, desirably for 30 minutes, and may be stored indefinitely. The complexes are premixed in serum free media. The nucleic acid and the transfecting reagent may be mixed either through direct addition or through a slow mixing process, such as across a dialyzing membrane or through the use of a microporous particle or a device that brings the two solutions together at a slow rate and at low concentrations. In some instances, the two interacting components are mixed at low concentrations, and the final complex is concentrated using a diafilteration or any other concentrating device. Alternatively, if the complexes are formed at high concentrations of either or both of the interacting components, the complexes may be diluted to form an ideal transfection mixture.

Transfection protocol and analysis of dsRNA stress response

[0181] Complexes are added to cells that are ~60 -80% confluent in serum free media. The complexes are incubated for various times (*e.g.*, 10 minutes to 24 hours) with the cells at 37°C and diluted with serum containing media or washed and replated in serum free media. The cells are monitored for toxicity and analyzed at various times for signs of dsRNA response (*e.g.*, TUNNEL assay to detect nicked DNA, phosphorylation of EIF2alpah, induction and activation of 2'5' OAS, or interferon-alpha and -beta). Transfection conditions that result in less than 50%, 25%, 10%, or 1% cytotoxicity or that result in a less than 20, 10, 5, 2, or 1.5-fold induction of a stress response are analyzed to determine if PTGS was efficiently induced.

Determining induction of PTGS

[0182] PSA protein levels are determined in cell culture media using standard methods. The data is normalized to the number of live cells in culture to determine the concentrations required to induce PTGS.

Results

[0183] Using the above method, cationic lipid complexes of dsRNA induced toxicity at certain ranges. With lipofectamine as the cationic lipid, positive to negative charge ratios greater than 10 did not produce any detectable toxicity at any of the concentrations of dsRNA tested and induced a high level of PTGS, resulting in highly decreased levels of PSA in the culture medium. The RNA concentration ranges tested were 1 pg to 100 ng with a constant amount of lipofectamine (10 uL of a 2 mg/mL solution from GIBCO-BRL Life Technologies,

Bethesda, MD).

Example 13: Method to avoid dsRNA mediated activation of the RNA stress response pathway

[0184] One or more components of the RNA stress response pathway can be mutated or inactivated to avoid induction/activation of the component(s) by dsRNA that is delivered to the cell or animal for the purpose of inducing PTGS. These components, such as those illustrated in Fig. 4, can be knocked out singularly or in combination.

[0185] Various standard methods can be used to knockout components of the RNA stress response pathway, such as PKR, human beta interferon Accession No. M25460), and/or 2'5'OAS (Accession No. NM_003733). Alternatively or additionally, one or more interferon response element (IRE) sequences can be mutated or deleted using a knockout construct designed based on the IRE consensus sequence (Ghislain, *et al.*, J Interferon Cytokine Res. 2001 Jun 21(6):379-88.), and/or one or more transcription factors that bind IRE sequences, such as STAT1 (Accession number XM_010893), can be mutated or deleted. These methods include the use of antisense DNA/RNA, ribozymes, or targeted gene knockout technology mediated by homologous recombination. One skilled in the art is able to design the appropriate antisense sequences, ribozymes, and vectors for targeted knockouts. For example, targeted knockouts may be prepared by any of the following standard methods: Shibata *et al.*, Proc Natl Acad Sci U S A. 2001 Jul 17;98(15):8425-32. Review., Muyrers *et al.*, Trends Biochem Sci. 2001 May;26(5):325-31., Paul *et al.*, Mutat Res. 2001 Jun 5;486(1):11-9., Shcherbakova *et al.*, Mutat Res. 2000 Feb 16;459(1):65-71., Lantsov. Ideal gene therapy: approaches and prospects Mol Biol (Mosk). 1994 May-Jun;28(3):485-95., in Gene Transfer and Expression - A Laboratory Manual editor: Michael Kriegler, Publisher- WH Freeman & Co, New York, NY, pages 56 -60, 1990).

[0186] Knockout cells can be readily identified either through the use of an antibiotic resistance marker which when transferred to the chromosome confers resistance to the cell or through the use of dsRNA itself. In particular, dsRNA (*e.g.*, a high concentration of dsRNA) Induces apoptosis in wild-type cells while mutant cells survive dsRNA treatment because they cannot mount a stress response. Yet another approach involves performing the dsRNA-induced PTGS experiment in the presence of large concentrations of IRE (dsDNA) oligo, which is expected to titrate activated STAT proteins. These oligos can be delivered intracerllularly using transfecting agents or electroporation.

[0187] In another method of preventing the interferon response, cells (*e.g.*, RD cells) are transfected with a T7 RNA polymerase expression vector and a T7 dsRNA expression vector encoding dsRNA homologous to the human protein kinase PKR cDNA (accession number M35663) or homologous to the coding sequence of any other component in the RNA stress response pathway. In one particular example, dsRNA corresponding to nucleotides 190 -2000 is encoded by the T7dsRNA expression vector. The expression vectors are similar to those described in Example 10 and shown in Fig. 2, except that the dsRNA encoding sequence is derived from the human protein kinase PKR cDNA. Transfection in RD cells is performed as described in Example 10. Within 2-5 days post-transfection, the cells are functionally PKR negative.

Other Embodiments

[0188] From the foregoing description, it will be apparent that variations and modifications may be made to the invention described herein to adopt it to various usages and conditions. Such embodiments are also within the scope of the following claims.

[0189] All publication, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

Claims

1. A method for identifying a nucleic acid that modulates the function of a vertebrate cell, said method comprising the steps of:

   (a) transforming a population of vertebrate cells with a double stranded RNA expression library, wherein at least two cells of said population of cells are each transformed with a different nucleic acid from said double stranded RNA expression library, wherein said transformed nucleic acid is capable of forming double stranded RNA, and wherein said transformation and said formation of double stranded RNA are carried out under conditions that inhibit or prevent an interferon response or a double stranded RNA stress response;

   (b) selecting a vertebrate cell in which said nucleic acid is expressed in said cell; and

   (c) assaying for a modulation in the function of said cell, wherein said modulation identifies a nucleic acid that modulates the function of said vertebrate cell.

2. A method for identifying a nucleic acid that modulates expression of a target nucleic acid in a vertebrate cell, said method comprising the steps of:

(a) transforming a population of vertebrate cells with a double stranded RNA expression library, wherein at least two cells of said population of cells are each transformed with a different nucleic acid from said double stranded RNA expression library, wherein said transformed nucleic acid is capable of forming double stranded RNA, and wherein said transformation and said formation of double stranded RNA are carried out under conditions that inhibit or prevent an interferon response or a double stranded RNA stress response;

(b) selecting a vertebrate cell in which said nucleic acid is expressed in said cell; and

(c) assaying for a modulation in the expression of a target nucleic acid in said cell, wherein said modulation identifies a nucleic acid that modulates expression of said target nucleic acid.

3. A method for identifying a nucleic acid that modulates the biological activity of a target polypeptide in a vertebrate cell, said method comprising the steps of:

(a) transforming a population of vertebrate cells with a double stranded RNA expression library, wherein at least two cells of said population of cells are each transformed with a different nucleic acid from said double stranded RNA expression library, wherein said transformed nucleic acid is capable of forming double stranded RNA, and wherein said transformation and said formation of double stranded RNA are carried out under conditions that inhibit or prevent an interferon response or a double stranded RNA stress response;

(b) selecting for a vertebrate cell in which said nucleic acid is expressed in said cell; and

(c) assaying for a modulation in the biological activity of a target polypeptide in said cell, wherein said modulation identifies a nucleic acid that modulates the biological activity of said target polypeptide.

4. The method of claim 1, wherein at most one nucleic acid is stably integrated into a chromosome of each cell.

5. The method of claim 1, said method further comprising:

(d) identifying said nucleic acid by amplifying said nucleic acid and sequencing said amplified nucleic acid.

6. The method of claim 1, wherein said double stranded RNA expression library comprises cDNAs derived from said cells.

7. The method of claim 1, wherein said double stranded RNA expression library comprises randomized nucleic acids.

8. The method of claim 1, wherein said double stranded RNA expression library is a nuclear double stranded RNA expression library.

9. The method of claim 1, wherein said double stranded RNA expression library is a cytoplasmic double stranded RNA expression library.

10. The method of claim 1, wherein said cell is a mammalian cell.

11. The method of claim 10, wherein said cell is a human cell.

12. The method of claim 10, wherein said cell is selected from the group consisting of a cancer cell, a cell of the immune system, a stem cell, a neuronal cell, a muscle cell, and an adipocyte.

13. The method of claim 1, wherein said nucleic acid is contained in a vector.

14. The method of claim 13, wherein said vector comprises an RNA polymerase II promoter, an RNA polymerase I promoter, an RNA polymerase III promoter, or a mitochondrial promoter.

15. The method of claim 13, wherein the sense strand and the anti-sense strand of said double stranded RNA are transcribed from the same nucleic acid using two convergent promoters.

16. The method of claim 13, wherein said nucleic acid comprises an inverted repeat such that upon transcription said nucleic acid forms a double stranded RNA.

17. The method of claim 1, wherein said assaying comprises measuring an event selected from the group consisting of cell motility, apoptosis, cell growth, cell invasion, vascularization, cell cycle events, cell differentiation, cell dedifferentiation, neuronal cell regeneration, and the ability of a cell to support viral replication.

18. The method of claim 1, wherein said double stranded RNA is between 5 and 100 nucleotides in length, inclusive.

19. The method of claim 1, wherein said double strand-

ed RNA is at least 100 nucleotides in length.

20. The method of claim 19, wherein said double stranded RNA is at least 250 nucleotides in length.

21. The method of claim 20, wherein said double stranded RNA is at least 500 nucleotides in length.

22. The method of claim 21, wherein said double stranded RNA is at least 1000 nucleotides in length.

23. A method for identifying a nucleic acid that modulates the function of a vertebrate cell, said method comprising the steps of:

(a) transforming a population of vertebrate cells with a double stranded RNA under conditions that inhibit or prevent an interferon response or a double stranded RNA stress response;

(b) selecting a vertebrate cell in which said double stranded RNA is expressed; and

(c) assaying for a modulation in the function of said cell, wherein said modulation identifies a nucleic acid that modulates the function of said cell.

24. A method for identifying a nucleic acid that modulates expression of a target nucleic acid in a vertebrate cell, said method comprising the steps of:

(a) transforming a population of vertebrate cells with a double stranded RNA under conditions that inhibit or prevent an interferon response or a double stranded RNA stress response;

(b) selecting a vertebrate cell in which said double stranded RNA is expressed; and

(c) assaying for a modulation in the expression of a target nucleic acid in said cell, wherein said modulation identifies a nucleic acid that modulates expression of said target nucleic acid.

25. A method for identifying a nucleic acid that modulates the biological activity of a target polypeptide in a vertebrate cell, said method comprising the steps of:

(a) transforming a population of vertebrate cells with a double stranded RNA under conditions that inhibit or prevent an interferon response or a double stranded RNA stress response;

(b) selecting a vertebrate cell in which said double stranded RNA is expressed; and

(c) assaying for a modulation in the biological activity of a target polypeptide in said cell, wherein said modulation identifies a nucleic acid that modulates the biological activity of said target polypeptide.

26. The method of claim 23, said method further comprising:

(d) identifying said nucleic acid by amplifying said nucleic acid and sequencing said amplified nucleic acid.

27. The method of claim 23, wherein said double stranded RNA is derived from said cells.

28. The method of claim 23, wherein said double stranded RNA is between 5 and 100 nucleotides in length, inclusive.

29. The method of claim 23, wherein said double stranded RNA is at least 100 nucleotides in length.

30. The method of claim 29, wherein said double stranded RNA is at least 250 nucleotides in length.

31. The method of claim 30, wherein said double stranded RNA is at least 500 nucleotides in length.

32. The method of claim 31, wherein said double stranded RNA is at least 1000 nucleotides in length.

33. The method of claim 23, wherein each cell is transformed with at most one double stranded RNA.

34. The method of claim 23, wherein said cell is a mammalian cell.

35. The method of claim 34, wherein said cell is a human cell.

36. The method of claim 23, wherein said population is transformed with at least 100 different double stranded RNAs.

37. The method of claim 1 or 23, said method further comprising:

(d) identifying said nucleic acid that modulates the function of said cell.

38. The method of claim 2 or 24, said method further comprising:

(d) identifying said nucleic acid that modulates expression of said target nucleic acid.

39. The method of claim 3 or 25, said method further

comprising:

(d) identifying said nucleic acid that modulates the biological activity of said target polypeptide.

Fig. 1

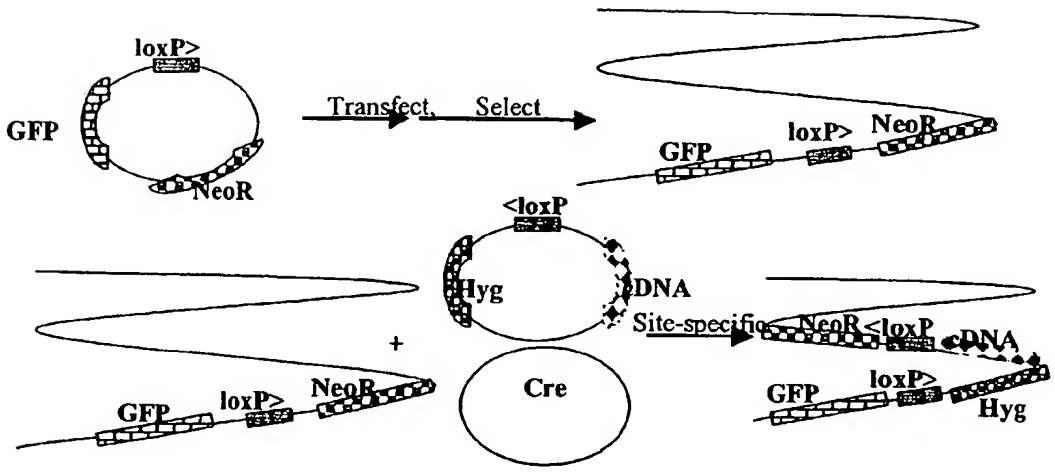

loxP>
GFP
NeoR
Transfect, Select
GFP
loxP>
NeoR
<loxP
Hyg
DNA
Site-specific
NeoR
<loxP
DNA
+
Cre
GFP
loxP>
NeoR
GFP
loxP>
Hyg

Figure 2

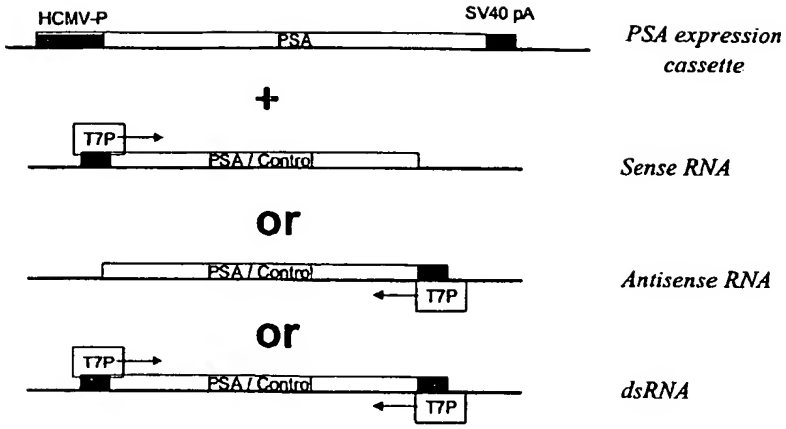


Figure 3

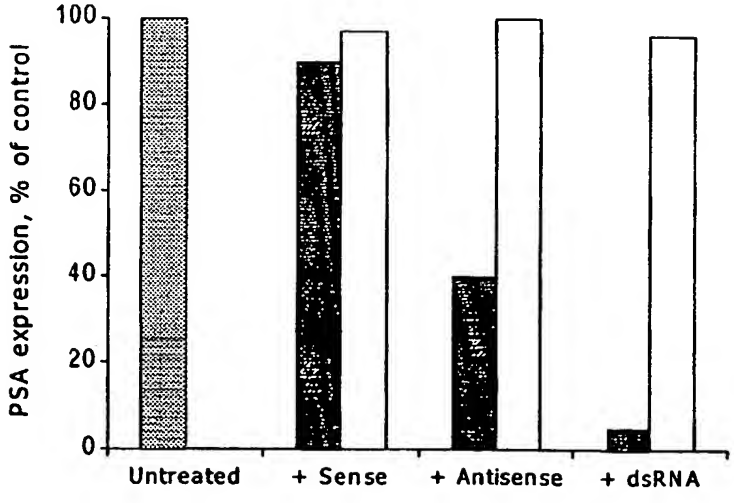
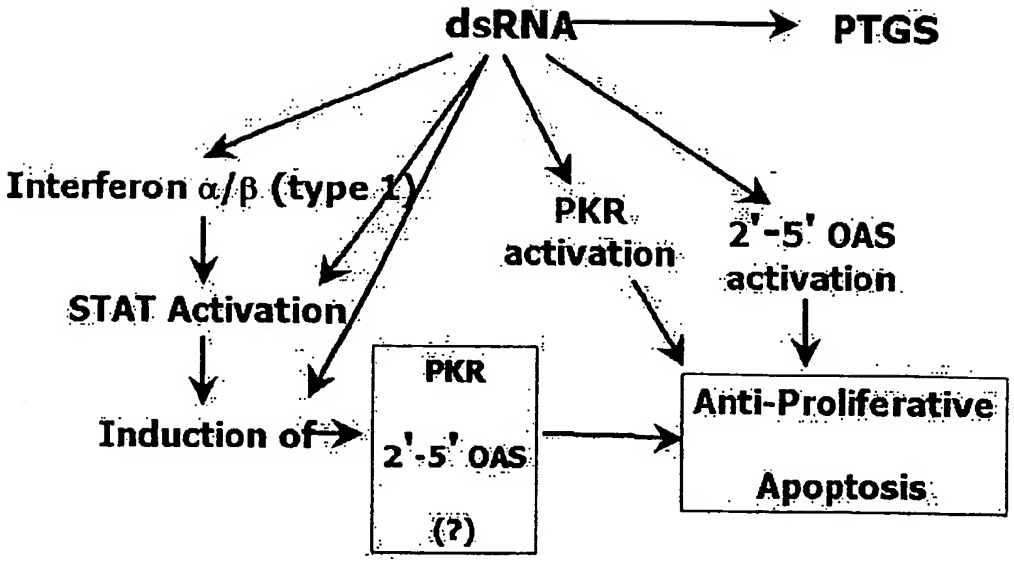


Figure 4


dsRNA
PTGS
Interferon α/β (type 1)
PKR activation
2'-5' OAS activation
STAT Activation
Induction of
PKR
2'-5' OAS
(?)
Anti-Proliferative
Apoptosis

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 229 134 A3

(12) EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
28.01.2004 Bulletin 2004/05

(43) Date of publication A2:
07.08.2002 Bulletin 2002/32

(21) Application number: **02250681.0**

(22) Date of filing: **31.01.2002**

(51) Int Cl.[7]: **C12Q 1/68**, C12N 15/10, G01N 33/50

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE TR
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **31.01.2001 US 265805 P**
26.10.2001 US 339260 P

(71) Applicants:
- **Nucleonics, Inc**
  Malvern, Pennsylvania 19355 (US)
- **Message Pharmaceuticals, Inc.**
  Malvern, PA 19355 (US)

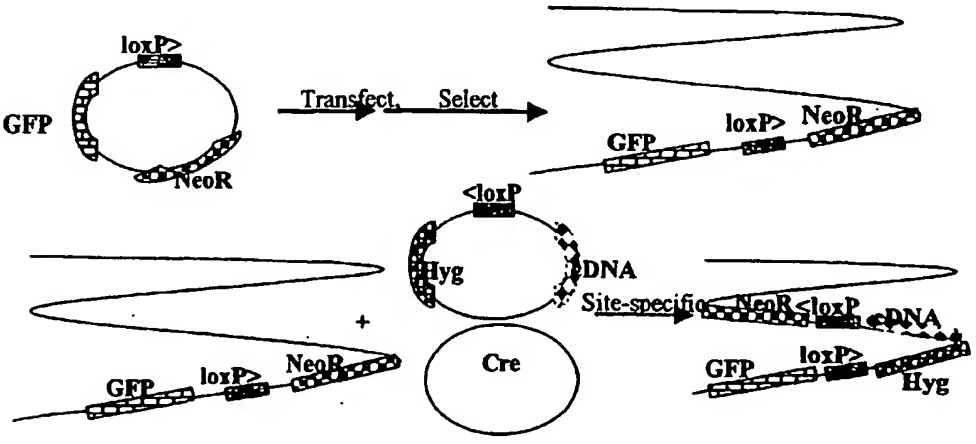
(72) Inventors:
- **Giordano, Tony**
  Pheonixville, PA 19460 (US)
- **Pachuk, Catherine**
  Lansdale, PA 19446 (US)
- **Satishchandran, Chandrasekhar**
  Lansdale, PA 19446 (US)

(74) Representative: **Davies, Jonathan Mark**
Reddie & Grose
16 Theobalds Road
London WC1X 8PL (GB)

(54) Use of post-transcriptional gene silencing for identifying nucleic acid sequences that modulate the function of a cell

(57) Described herein are methods for identifying nucleic acid sequences that modulate the function of a cell, the expression of a gene in a cell, or the biological activity of a target polypeptide in a cell. The methods involve the use of double stranded RNA expression libraries, double stranded RNA molecules, and post-transcriptional gene silencing techniques.

Fig. 1



EP 1 229 134 A3

Printed by Jouve, 75001 PARIS (FR)

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 02 25 0681

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
X	UI-TEI K ET AL: "Sensitive assay of RNA interference in Drosophila and Chinese hamster cultured cells using firefly luciferase gene as target" FEBS LETTERS, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL, vol. 479, no. 3, 18 August 2000 (2000-08-18), pages 79-82, XP004337469 ISSN: 0014-5793 * abstract * * page 81, column 1, paragraph 2 - page 82, column 1, paragraph 4; figures 1,3 * ---	1-39	C12Q1/68 C12N15/10 G01N33/50
X	WIANNY F ET AL: "Specific interference with gene function by double-stranded RNA in early mouse development" NATURE CELL BIOLOGY, MACMILLAN PUBLISHERS, GB, vol. 2, no. 2, February 2000 (2000-02), pages 70-75, XP002138445 ISSN: 1465-7392 * abstract * * page 73, column 2, paragraph 1 - page 74, column 1, paragraph 3; figures 2,3,5,6 * --- -/--	23-39	TECHNICAL FIELDS SEARCHED (Int.Cl.7) C12N C12Q

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
MUNICH	9 December 2003	Bradbrook, D

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 02 25 0681

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
X	SVOBODA P ET AL: "SELECTIVE REDUCTION OF DORMANT MATERNAL MRNAS IN MOUSE OOCYTES BY RNA INTERFERENCE" DEVELOPMENT, COMPANY OF BIOLOGISTS, CAMBRIDGE,, GB, vol. 127, no. 127, 2000, pages 4147-4156, XP001064763 ISSN: 0950-1991 * abstract * * page 4152, column 1, paragraph 3 - page 4153, column 2, paragraph 1 * * page 4154, column 2, paragraphs 1,2; figures 2,3,5,6 * ---	23-39	
X	WARGELIUS A ET AL: "DOUBLE-STRANDED RNA INDUCES SPECIFIC DEVELOPMENTAL DEFECTS IN ZEBRAFISH EMBRYOS" BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, ACADEMIC PRESS INC. ORLANDO, FL, US, vol. 263, no. 1, 16 September 1999 (1999-09-16), pages 156-161, XP000872838 ISSN: 0006-291X * abstract; figure 2; table 1 * ---	23-39	TECHNICAL FIELDS SEARCHED (Int.Cl.7)
X	ELBASHIR SAYDA M ET AL: "RNA interference is mediated by 21- and 22-nucleotide RNAs" GENES AND DEVELOPMENT, COLD SPRING HARBOR LABORATORY PRESS, NEW YORK, US, vol. 15, no. 2, 15 January 2001 (2001-01-15), pages 188-200, XP002204651 ISSN: 0890-9369 * abstract; figures 1,5 * --- -/--	23-39	

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
MUNICH	9 December 2003	Bradbrook, D

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 02 25 0681

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
X	WO 00 63364 A (AMERICAN HOME PROD ;PACHUK CATHERINE (US); SATISHCHANDRAN C (US)) 26 October 2000 (2000-10-26) * page 8, line 6 - page 9, line 28 * * page 17, line 12 - page 20, line 16; examples 1-4 * ---	1-39	
X	WO 99 32619 A (CARNEGIE INST OF WASHINGTON ;MONTGOMERY MARY K (US); FIRE ANDREW ()) 1 July 1999 (1999-07-01) * page 6, line 1 - page 8, line 2 * * page 9, line 15 - page 20, line 15; claims 1,21 * ---	1-39	
A	BOSHER J M ET AL: "RNA INTERFERENCE: GENETIC WAND AND GENETIC WATCHDOG" NATURE CELL BIOLOGY, MACMILLAN PUBLISHERS, GB, vol. 2, no. 2, February 2000 (2000-02), pages E31-E36, XP001106960 ISSN: 1465-7392 * the whole document * ---	1-39	TECHNICAL FIELDS SEARCHED (Int.Cl.7)
P,X	ELBASHIR SAYDA M ET AL: "Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells" NATURE, MACMILLAN JOURNALS LTD. LONDON, GB, vol. 411, no. 6836, 24 May 2001 (2001-05-24), pages 494-498, XP002206451 ISSN: 0028-0836 * the whole document * ---	23-39	
P,A	WO 01 75164 A (MAX PLANCK GESELLSCHAFT ;TUSCHL THOMAS (DE); MASSACHUSETTS INST TE) 11 October 2001 (2001-10-11) * the whole document * --- -/--		

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
MUNICH	9 December 2003	Bradbrook, D

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 02 25 0681

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
E	WO 02 44321 A (MAX PLANCK GESELLSCHAFT ;TUSCHL THOMAS (DE); ELBASHIR SAYDA (DE);) 6 June 2002 (2002-06-06) * the whole document *	1-39	

TECHNICAL FIELDS SEARCHED (Int.Cl.7)

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
MUNICH	9 December 2003	Bradbrook, D

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 02 25 0681

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

09-12-2003

Patent document cited in search report		Publication date		Patent family member(s)	Publication date
WO 0063364	A	26-10-2000	AU	4472100 A	02-11-2000
			BR	0009884 A	08-01-2002
			CA	2370628 A1	26-10-2000
			CN	1375004 T	16-10-2002
			EP	1171586 A2	16-01-2002
			JP	2002542263 T	10-12-2002
			WO	0063364 A2	26-10-2000
WO 9932619	A	01-07-1999	US	6506559 B1	14-01-2003
			AU	743798 B2	07-02-2002
			AU	1938099 A	12-07-1999
			CA	2311999 A1	01-07-1999
			EP	1042462 A1	11-10-2000
			JP	2002516062 T	04-06-2002
			WO	9932619 A1	01-07-1999
			US	2003056235 A1	20-03-2003
			US	2003051263 A1	13-03-2003
			US	2003055020 A1	20-03-2003
WO 0175164	A	11-10-2001	AU	3574402 A	11-06-2002
			AU	4962201 A	15-10-2001
			CA	2404890 A1	11-10-2001
			CA	2429814 A1	06-06-2002
			CZ	20031839 A3	15-10-2003
			WO	0244321 A2	06-06-2002
			EP	1309726 A2	14-05-2003
			HU	0302557 A2	28-10-2003
			JP	2003529374 T	07-10-2003
			NO	20032464 A	21-07-2003
			WO	0175164 A2	11-10-2001
			US	2003108923 A1	12-06-2003
			US	2002086356 A1	04-07-2002
WO 0244321	A	06-06-2002	AU	3574402 A	11-06-2002
			AU	4962201 A	15-10-2001
			CA	2404890 A1	11-10-2001
			CA	2429814 A1	06-06-2002
			CZ	20031839 A3	15-10-2003
			WO	0244321 A2	06-06-2002
			EP	1309726 A2	14-05-2003
			HU	0302557 A2	28-10-2003
			JP	2003529374 T	07-10-2003
			NO	20032464 A	21-07-2003
			WO	0175164 A2	11-10-2001
			US	2003108923 A1	12-06-2003
			US	2002086356 A1	04-07-2002

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

RELATED PROCEEDINGS APPENDIX

There are no related proceedings.